GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
(without alignments)

Run on: December 23, 2005, 18:47:32 ; Search time 4894 Seconds
4169.759 Million cell updates/sec

Title: US-10-019-931-3
Perfect score: 1898
Sequence: 1 MATSSMKSIPMAIPSFSMCH..........RLSALVDDKLGGSGLTPRLS 359

Scoring table: BLOSUM62
                 Xgapop 10.0 ; Xgapext 0.5
                 Ygapop 10.0 ; Ygapext 0.5
                 Pgapop 6.0 ; Pgapext 7.0
                 Delop   6.0 ; Delext 7.0

Searched: 5883141 seqs, 2842172565 residues

Total number of hits satisfying chosen parameters: 11766282

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1p
-Q=/cgn2_1/USPTO_spool_p/US10019931/runat_16122005_165218_18438/app_query.fasta_1.519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10019931@CGN_1_1_7415 @runat_16122005_165218_18438 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_ph:*
7: gb_pat:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1898	100.0	1077	6	AX080458	AX080458 Sequence
2	1898	100.0	1080	15	BT008847	BT008847 Arabidops
3	1898	100.0	1273	6	CQ832234	CQ832234 Sequence
4	1898	100.0	1273	6	CQ801446	CQ801446 Sequence
5	1898	100.0	1273	15	AY084999	AY084999 Arabidops
6	1898	100.0	1347	15	AY099809	AY099809 Arabidops
7	1898	100.0	86380	15	AB010697	AB010697 Arabidops
8	1892	99.7	1270	15	AY054219	AY054219 Arabidops
9	1582	83.4	1236	15	AK172202	AK172202 Arabidops
10	1553	81.8	1041	6	AX080460	AX080460 Sequence
11	1553	81.8	1236	15	BT006241	BT006241 Arabidops
12	1519	80.0	1044	15	CQ832235	CQ832235 Sequence
13	1519	80.0	1023	6	CQ801447	CQ801447 Sequence
14 c	1076.5	56.7	115071	14	AC135566	AC135566 Medicago
15 c	1024	54.0	100108	14	AP008111	AP008111 Lotus cor
16	937	49.4	115362	15	AP004534	AP004534 Lotus cor
17 c	905.5	47.7	116174	15	AC149197	AC149197 Medicago
18 c	905.5	47.7	116174	15	AC148349	AC148349 Medicago
19 c	899.5	47.1	112154	15	AC148395	AC148395 Medicago
20 c	876	47.4	108265	15	AC146342	AC146342 Medicago
21 c	861.5	46.2	119355	14	AC119410	AC119410 Medicago
22 c	861.5	45.4	145905	15	AC144476	AC144476 Medicago
23 c	861	45.4	111277	15	AC147431	AC147431 Medicago
24	861	45.4	111277	15	AC150505	AC150505 Medicago
25 c	858.5	45.4	138586	15	AC144726	AC144726 Medicago
26 c	858.5	45.2	111277	15	AC150505	AC150505 Medicago
27	858	45.2	111277	15	AC150505	AC150505 Medicago
28 c	850.5	45.2	138586	14	AC144726	AC144726 Medicago
29	848.5	45.2	93941	15	AC150505	AC150505 Medicago
30 c	783.5	44.8	136553	15	AC144726	AC144726 Medicago
31 c	776	44.7	112917	14	AC149307	AC149307 Medicago
32 c	776	41.3	180062	14	AC123976	AC123976 Medicago
33	776	40.9	1090	6	AC151963	AC151963 Glycine m
34	776	40.9	1154	15	BT002108	BT002108 Arabidops
35	770.5	40.9	1216	15	AY087921	AY087921 Arabidops
36	770.5	40.6	996	6	AY092961	AY092961 Arabidops
37 c	757	40.6	87543	15	AC006836	AC006836 Arabidops
38	757	39.9	996	6	AX506332	AX506332 Sequence
39	757	39.9	980	6	AF325065	AF325065 Arabidops
40	757	39.9	996	6	AX651508	AX651508 Sequence
41	757	39.9	981	6	AX412507	AX412507 Sequence
42	757	39.9	981	6	AX412570	AX412570 Sequence
43	757	39.9	981	6	AX507354	AX507354 Sequence
44	742	39.9	1179	15	AX589857	AX589857 Sequence
45	742	39.1	996	6	AY113050	AY113050 Arabidops

AF375458 Arabidops
BT015380 Arabidops

ALIGNMENTS

RESULT 1
AX080458

```
LOCUS       AX080458                1077 bp    DNA     linear   PAT 22-FEB-2001
DEFINITION  Sequence 1 from Patent WO0102589.
ACCESSION   AX080458
VERSION     AX080458.1  GI:13159885
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (residues 1 to 1077)
  AUTHORS   Varin,L. and Gidda,S.
  TITLE     Methods, compositions and genetic sequences for modulating
            flowering in plants, and plants genetically modified to flower
            early and tardily
  JOURNAL   Patent: WO 0102589-A 1 11-JAN-2001;
            Varin, Luc (CA) ; Gidda, Satinder (CA)
FEATURES             Location/Qualifiers
     source          1..1077
                     /organism="Arabidopsis thaliana"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3702"
ORIGIN
Alignment Scores:
```

```
US-10-019-931-3 (1-359) x AX080458 (1-1077)

Pred. No.:        1             Length:            1077
Score:            1.42e-156     Matches:           359
Percent Similarity: 100.00%     Conservative:      0
Best Local Similarity: 100.00%  Mismatches:        0
Query Match:      6             Indels:            0
DB:                             Gaps:              0
```

```
Qy    1  MetAlaThrSerSerMetLysValSerIleProMetAlaIleProSerPheSerMetCysHis  20
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  ATGGCTACCTCAAGCATGAAAGTGAGCATTCCAATGGCCATCCCAAGTTTCTCCATGTGTCAC  60

Qy   21  LysLeuGluLeuLeuLysGlyLysThrArgAspValProLysAlaGluGluAspGlu  40
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  AAGCTCGAGCTCCTTAAAGGAAAGAACCCGACAGTTCCGAAAGCCGAAGAAGATGAA  120

Qy   41  GlyLeuSerCysGluPheGlnGluMetLeuAspSerLeuProLysGlyTrpArg  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GGGCTAAGCTGCGAGTTCCAAGAGATGTTGGATTCTCTTCCTAAGGAGGATGGAGA  180

Qy   61  ThrArgTyrLeuTyrLeuPheGlnGlyPheTrpCysGlnAlaAlaLysGlyAlaAlaIle  80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  ACTCGTTACCTTACCTATTCCAACGGTTTTGGTGCCAAGCCAAAGAGATTCAAGGCCATC  240

Qy   81  MetSerPheGlnLysHisPheGlnSerLeuGluAsnAspValValLeuAlaThrIlePro  100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  ATGTCTTTCAAAAAACATTTCCAATCCCTCGAAAACGACGTCGTTCTCGCCACCATACCT  300

Qy  101  LysSerGlyThrThrTrpLeuTyrLeuAlaLeuThrPheThrIleLeuAsnArgHisArgPhe  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  AAAATCCGGTACAACCTGGCTAAAAGCTTTAACTTACACAATCCTTAACCGTACCGGGTTT  360

Qy  121  AspProAlaAlaSerThrArgAsnHisProLeuPheThrSerAsnProHisAspLeuVal  140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  GATCCGGTTGCCTCCGAGTACCAAACCACCCTCTTTCACTTCCAACCCTCATGACCTTGTA  420

Qy  141  ProPhePheGluPheTyrLysLeuTyrAlaAsnGlyAspValProAspLeuSerGlyLeuAla  160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  CCTTTCTTCGAGTTCTACAAGCTTTACGCCAACGGAGATGTCCCAGACCTCTCGGGTCTAGCC  480

Qy  161  SerProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGlyLysThrIleGluLys  180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  AGTCCAAGAACGTTCGCCACCCACTTACCGTTCGGTTCCTTAAAGGAAGGAAATCAGAGAAA  540

Qy  181  ProGlyValLysValAlaTyrLeuCysArgAsnProPheAspThrPhePheSerSerTrp  200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  CCCGGTGTGAAGGTCGTGTACTTGTGCCGGAACCCGTTTGACACNATTCATCCTCGTGG  600

Qy  201  HisTyrThrAsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaAlaPhe  220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  CATTACACCAACATCAAAATCCGAGTCGACCCCAGTCTTGCTCGACCAAGCTGTTT  660

Qy  221  AspLeuTyrCysArgGlyValIleGlyTyrProGlyProPheTrpGluHisMetLeuGlyTyr  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  GATCTGTATTGCCGGGAGGTGATCGGGTTTGGGACACATGTTGGGATAC  720

Qy  241  TrpArgGluSerLeuLysArgProGluLysProGluLeuLys  260
         ||||||||||||||||||||||||||||||||||||||||||
Db  721  TGGAGAGAGAGCTTGAAGAGACCAGAGAAAGCTTCTTTTAAGGTACGGACGGATCTCAAA  780

Qy  261  AspAspIleGluThrPheLeuLysGlyLeuProPheThrGlu  280
         ||||||||||||||||||||||||||||||||||||||||||
Db  781  GACCGACATCGAGACCTTCCTGAAGGGTTTGCCGACTTTCTTTAGAGCCTTCCTTTCACCGAA  840

Qy  281  GluGluLeuArgArgGlyValValAlaIleLeuLeuCysSerSerPheGluAsnLeu  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  GAAGAGGAACGAAAGGGAGTTGTGAAGGCTATCCTCGGAGCTTCGTAGCTTCGAGAATCTCG  900

Qy  301  LysLysLeuGluValAlaSerLysSerAsnLysLysSerAsnPheGluAsnArgPheLeu  320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  AAGAAGTTGGAGGTCGAAAGCAAGTCAAACAAGAAGTCGAATCAAGAACTTGAGAATCGATTCTTG  960

Qy  321  PheArgLysGlyGlyValSerAspTrpValAlaAsnTyrLeuSerProSerGlnValGlyArg  340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  TTTCGGAAAGGAGGAGTGAGTGATTGGGTTGCAAACTATTTGTCACCTTCACAAGTTGGAAGA  1020

Qy  341  LeuSerAlaLeuValAspAspGlyLeuGlyGlySerGlyLeuThrPheArgLeuSer  359
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  TTGTCAGCCTTAGTTGGTGATGACAAGTTAGGTGGATCCGGTCTCACTTTCAGGTTGAGC  1077
```

```
RESULT 2
BT008847
LOCUS       BT008847              1080 bp    mRNA    linear   PLN 13-JUN-2003
DEFINITION  Arabidopsis thaliana At5g07010 mRNA, complete cds.
ACCESSION   BT008847
VERSION     BT008847.1  GI:31711859
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1080)
  AUTHORS   Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P.,
            Dale,J.M., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A.,
            Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M.,
            Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L.,
            Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C.,
            Wu,H.C., Yamada,K., Yu,G., Shinozaki,K., Davis,R.W., Theologis,A.
            and Ecker,J.R.
  TITLE     Arabidopsis ORF clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1080)
  AUTHORS   Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P.,
            Dale,J.M., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A.,
            Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M.,
            Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L.,
            Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C.,
            Wu,H.C., Yamada,K., Yu,G., Shinozaki,K., Davis,R.W., Theologis,A.
            and Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUN-2003) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA = 'RIKEN
            Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.
            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim,C.J.,
            Chen,H., Cheuk,R., Shinn,P., Bowser,L., Chan,M.M., Chang,C.M.,
            Dale,J.M. Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M, Lin,J.,
            Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L.,
            Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K.,
            Yu,G., Davis,R.W., Theologis,A. and Ecker,J.R.
            Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1..1080
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="U25564"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51"
     CDS             1..1080
                     /note="steroid sulfotransferase-like protein"
                     /codon_start=1
                     /product="At5g07010"
```

ORIGIN

```
/protein_id="AAP68286.1"
/db_xref="GI:31711860"
/translation="MATSSMKSIPMAIPSFSMCHKLELLKEGKTRDVPKAEEDEGLSC
EFQEMLDSLPKERGWRTRYLYLFQGFWCQAKEIOAIMSFQKHFQSLENGDVULAIIPKS
GTTWKALFTILNRHRFDPVASSTNHPLTYSNPHDLVFFEYKLYANGUVPDLSGLA
SPRTFATHLPPGSLKETIEKPGVKVVYLCNPFDTFISSWHYTNNIKSESVSPVLLDQ
AFDLYCRGVIGFGPPFWEHMLGYWRESLKRPEKVPFLRYEDLKDDIETNLKRLATFLBL
PFTEEERKGVKVKAIAELCSFENLKKLIEVNKSNKIKNFENRFLFRKGEVSDWVNYLS
PSQVERLSALVDDKLGGSGLTPRLS"
```

US-10-019-931-3 (1-359) x BT008847 (1-1080)

Alignment Scores:
```
Pred. No.:
Score:                    1.43e-156
                          1898.00
Percent Similarity:       100.00%
Best Local Similarity:    100.00%
Query Match:              15
DB:
```

```
Length:       1080
Matches:      359
Conservative: 0
Mismatches:   0
Indels:       0
Gaps:         0
```

```
Qy     1 MetAlaThrSerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHis 20
Db     1 ATGGCTACCTCAAGCATGAAGAGCCACTTCCAATGGCGATCCCAAGTTTCTCACGTGTCAC 60

Qy    21 LysGluLeuLeuLysGluGlyGlyLysThrArgAspValProLysAlaGluGluAspGlu 40
Db    61 AAGGCTCGAGCTCCTTAAAGAAGGCAAAACTCGGACGTCCCGAAAGCCGAAGAAGATGAA 120

Qy    41 GlyLeuSerCysGluPheGlnGluMetLeuAspSerLeuProLysGluArgGlyTrpArg 60
Db   121 GGGCTAAGCTGCGAGTTCCAAGAGATGTTGGATTCTCTTCCTAAGGAGAGAGGATGGAGA 180

Qy    61 ThrArgTyrLeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGluIleGlnAlaIle 80
Db   181 ACTCGTTACCTTTACCTATTCCAAGGGTTTTGGTGCCAAGCCAAAAGAGATTCAAGCCATC 240

Qy    81 MetSerPheGlnLysHisPheGlnSerLeuGluAsnGlyAspValValLeuAlaThrIlePro 100
Db   241 ATGTCTTTCCAAAAACATTTCCAAAGTCCCTGGAAAACGACGTCGTTCTCGCCACCATACCT 300

Qy   101 LysSerGlyThrThrTrpLysAlaLeuPheThrIleLeuAsnArgHisArgPhe 120
Db   301 AAATCCGGTACAACCTGGCCTAAAAGCTTAAACTTTCACCATCCTTAACGTCACCGGTTT 360

Qy   121 AspProValAlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAspLeuVal 140
Db   361 GATCCGGTTGCCTCGAGTACCAACCACCCTCTTTTCACTTCAAACCCTCATGACCTTGTA 420

Qy   141 PhePhePheGluTyrLysLeuTyrAlaAsnGlyAspValProAspLeuSerGlyLeuAla 160
Db   421 CCTTTCTTCTTCGAGTACAAGCTTTACGCCAACGGAGATGTTCCCGATCTCTCGGGTCTAGCC 480

Qy   161 SerProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluLys 180
Db   481 AGTCCAAGAACGTTCGCAACCCACTTACCGGTCCCTTAAAGGAAACGATCGAAAAG 540

Qy   181 ProGlyValLysValTyrLeuCysAsnProPheAspThrPheIleSerSerTrp 200
Db   541 CCCGGGTGTGAAGGTCGTGTACTTGTGCCGACACCGTTCACATTCATCTTCGTGG 600

Qy   201 HisTyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaPhe 220
Db   601 CATTACACCAACAACATCAAATCCGAGTCAGTGAGCCCAGTTCTTCTGGACCAAGCTTTT 660

Qy   221 AspLeuTyrCysArgGlyValIleGlyPheGlyProPheTrpGluHisMetLeuGlyTyr 240
Db   661 GATCTGTATTGCCGGGGAGTCATCGGGTTTGGCCCGTTTTGGGAACACATGTTGGGATAC 720

Qy   241 TrpArgGluSerLeuLysArgProGluLysValPhePheLeuArgTyrGluAspLeuLys 260
Db   721 TGGAGAGAGAGCTTGAAGAGACCAGAGAAAGTCTTCTTTTTAGGTACGAGGAGTCTCAAA 780
```

US-10-019-931-3.p2n.rge   Page 3

RESULT 3
```
LOCUS        CQ832234              1273 bp    DNA        linear    PAT 29-JUL-2004
DEFINITION   CQ832234 Sequence 1 from Patent WO2004057000.
ACCESSION    CQ832234
VERSION      CQ832234.1    GI:50831890
KEYWORDS     .
SOURCE       Arabidopsis thaliana (thale cress)
  ORGANISM   Arabidopsis thaliana
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
             rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1
  AUTHORS    Varin,L.
  TITLE      Methods and compositions for modulating tuber formation in
             tuber-producing plants
  JOURNAL    Patent: WO 200405700-A 1 08-JUL-2004;
             Florigene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1..1273
                     /organism="Arabidopsis thaliana"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3702"
```

ORIGIN

US-10-019-931-3 (1-359) x CQ832234 (1-1273)

Alignment Scores:
```
Pred. No.:
Score:                    1.72e-156
                          1898.00
Percent Similarity:       100.00%
Best Local Similarity:    100.00%
Query Match:              6
DB:
```

```
Length:       1273
Matches:      359
Conservative: 0
Mismatches:   0
Indels:       0
Gaps:         0
```

```
Qy     1 MetAlaThrSerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHis 20
Db    66 ATGGCTACCTCAAGCATGAAGAGCCACTTCCAATGGCGATCCCAAGTTTCTCACGTGTCAC 125

Qy    21 LysGluLeuLeuLysGluGlyGlyLysThrArgAspValProLysAlaGluGluAspGlu 40
Db   126 AAGGCTCGAGCTCCTTAAAGAAGGCAAAACTCGGACGTCCCGAAAGCCGAAGAAGATGAA 185

Qy    41 GlyLeuSerCysGluPheGlnGluMetLeuAspSerLeuProLysGluArgGlyTrpArg 60
Db   186 GGGCTAAGCTGCGAGTTCCAAGAGATGTTGGATTCTCTTCCTAAGGAGAGAGGATGGAGA 245

Qy    61 ThrArgTyrLeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGluIleGlnAlaIle 80
Db   246 ACTCGTTACCTTTACCTATTCCAAGGGTTTTGGTGCCAAGCCAAAAGAGATTCAAGCCATC 305

Qy    81 MetSerPheGlnLysHisPheGlnSerLeuGluAsnGlyAspValValLeuAlaThrIlePro 100
Db   306 ATGTCTTTCCAAAAACATTTCCAAAGTCCCTGGAAAACGACGTCGTTCTCGCCACCATACCT 365
```

Right column (continuation):

```
Qy   261 AspAspIleGluThrAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGlu 280
Db   781 GACGACATCGAGACCAACTTGAAGAGGCTTGCAACTTTCTTGAGCTTCCTTTCACCGGAA 840

Qy   281 GluGluGluArgGlyValLysValLysAlaIleAlaGluLeuCysSerPheGluAsnLeu 300
Db   841 GAAGAGGAACGAAAAGGCGTCAAGGTGAAGGCTATCGCCGAGCTGTGTTCGTTCGAGAATCTG 900

Qy   301 LysLysLeuGluValAsnLysSerAsnLysIleLysAsnPheGluAsnArgPheLeu 320
Db   901 AAGAAGTTGGAGGTGAACAAGTCAAACAAGATCAAGAACTTTGAGAATCGATTCTTG 960

Qy   321 PheArgLysGlyGluValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArg 340
Db   961 TTTCGGAAAGGAGAAGTGAGTGATTGGGTTAACTATTTGTCACCTTCACAAGTGGAAAGA 1020

Qy   341 LeuSerAlaLeuValAspAspLysLeuGlyGlySerGlyLeuThrProArgLeuSer 359
Db  1021 TTGTCAGCCTTAGTTGATGACAAGTTAGGTGGATCTGGTCTCACTTTCAGGTTGAGC 1077
```

```
FEATURES
  source
        Location/Qualifiers
        1..1273

JOURNAL     Florisys Inc. (CA); Institut fuer Pflanzenbiochemie (DE)
            Location/Qualifiers
            1..1273
            Patent: WO 200406235O-A 1 29-JUL-2004;
            Florisys Inc. (CA); Institut fuer Pflanzenbiochemie (DE)
TITLE       Methods and compositions for producing male sterile plants
AUTHORS     Maucher,H., Miersch,O., Wasternack,C. and Varin,L.
REFERENCE   1

KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
VERSION     CQ881446.1  GI:54034405
ACCESSION   CQ881446
DEFINITION  Sequence 1 from Patent WO2004062350.
LOCUS       CQ881446              1273 bp    DNA     linear   PAT 11-OCT-2004
```

```
RESULT 4
CQ881446
```

ORIGIN

```
Alignment Scores:
Pred. No.:
Score:                    1898.00    Matches:          359
Percent Similarity:       100.00%    Conservative:     0
Best Local Similarity:    100.00%    Mismatches:       0
Query Match:                         Indels:           0
DB:                                  Gaps:             0
                          100.00%

Length:         1273
Matches:        359

/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
```

US-10-019-931-3 (1-359) x CQ881446 (1-1273)

```
Qy      1  MetAlaThrSerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHis  20
Db     66  ATGGCTACCTCAAGCATGAAGAGCAATTCCAATGGCGGCATCCCAAGTTTCTCCATGTGTCAC 125

Qy     21  LysLeuGluLeuLeuLysLysGluGlyValLysThrArgAspValProLysGluGluGlu  40
Db    126  AAGCTCGAGCTCCTTAAAGAAGGGGAAAACGTCCGGGAAAGCCGAAGAGATGAA 185

Qy     41  GlyLeuSerCysGluPheGlnGlnMetLeuAspSerLeuProLysGluArgGlyTyrTrpArg 60
Db    186  GGGCTAAGCTGCGAGTTCCAAGAAGATGTTGGATTTCCAAGGAAGAGAGGAGA 245

Qy     61  ThrArgTyrLeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGluGlnAlaIle  80
Db    246  ACTCGTTACCTTTACCTATTCCAAGGGTTTGGGTGCCAAGCAAAGAGATTCAAGCCATC 305

Qy     81  MetSerPheGlnLysHisPheGlnSerLeuGluAsnAspValValLeuAlaThrIlePro 100
Db    306  ATGTCTTTCCAAAAACATTTCCAATCCCTCGAAAACGACGTCGTTCTCGCCACATACCT 365

Qy    101  LysSerGlyThrThrTrpTyrLeuTrpLeuAlaLeuThrArgHisArgPhe 120
Db    426  AAATCCGGTACAACCTGGTACCTACCGCTGGCTCTGACCAGACACCGGTTT 425

Qy    121  AspProValAlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAlaPheVal 140
Db    426  GATCCGGTTGCCTCGAGTACAACCACCCTCTTTCACTTCCAACCCTCATGACCTTGTA 485

Qy    141  ProPhePheGluTyrLysLeuTyrAlaAsnGlyAspValProAspLeuSerGlyLeuAla 160
Db    486  CCTTTCTTCGAGTACAAGCTTTACGCCAACGGAGATGTTCCCGAATCTCTCGGGTCTAGCC 545

Qy    161  SerProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluLys 180
Db    546  AGTCCAAGGACGTTCGCAACCCACTTACCGTTCGGTTCCCTAAAGGAAACGATCGAGAAA 605

Qy    181  ProGlyValLysValValTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrp 200
Db    666  CCTGGTGTGAAAGTCGTGTACTTGTGCCGGAACCCGTTTGACACATTCATCTCTTCGTGG 665

Qy    201  HisTyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaPhe 220
Db    726  CATTACACAAACAACATCAAATCCGAGTCAGTGTCACCAGTCCTTGCTCGACCAAGCTTTT 725

Qy    221  AspLeuTyrCysArgGlyValIleGlyPheGlyProTrpProGluTrpLeuAspLeuVal 240
Db    786  GATCTGTATTGCCGGGGAGTTATCGGTTTGGGCCGTTTGGCCGGAGTGGATCTTGTTGTA 785

Qy    241  TrpArgGluGlyGluAsnPheGluAsnLeuLysGlyValValAlaAlaIleLysAlaIle 260
Db    846  TGGAGAGAGGGCGAAAACTTTGAGAATCTCAAAGGAGTCGTTGCTGCAATTAAAGCCATC 845

Qy    261  AspAspIleGluThrAsnLeuLeuLysArgLeuAlaThrPheLeuGlyLeuProPheThrGlu 280
Db    906  GACGACATCGAGACCAACTTGCTGAAGCGTTTAGCAACTTTCCTTAGGACTCGAGGGATCTCAAA 905

Qy    281  GlyGlyGluGlyTyrValValLeuAlaGluLeuCysSerPheGluAsnLeu 300
Db    966  GGAGGAGAGGGCTACGTTGTGGGTTAACTTTGTCACCTTCACAAGTCGGAAAGA 965

Qy    301  LysGlyLysSerLeuGluLysValAsnLysSerAsnPheGluAsnPheLeuLysAspLeu 320
Db   1026  AAGAAGGTTGGAGAGGTTGAGAAGTTGGGGTTTAACTTTGTCAC 1025

Qy    321  PheArgGlyGlyGluValLeuAspTrpLeuAsnTyrLeuSerProSerGlnValGluArg 340
Db   1086  TTTCGGAAAGGAGAAGTGAGTGACTAGGTAGGTCTCACTTTCAGGTTGAGC 1085

Qy    341  LeuSerAlaLeuValAlaAspProLysGlyLeuGlySerGlyLeuThrPheArgLeuSer 359
Db   1142  TTGTCAGCCTTAGTGGATGACAAGTTAGGTCGATCTGGTCTCACTTTCAGGTTGAGC 1142
```

RESULT 5

```
LOCUS       AY084999                1273 bp    mRNA    linear   PLN 14-APR-2003
DEFINITION  Arabidopsis thaliana clone 124067 mRNA, complete sequence.
ACCESSION   AY084999
VERSION     AY084999.1  GI:21403709
KEYWORDS    FLI cDNA.
SOURCE      Arabidopsis thaliana
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1273)
  AUTHORS   Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
            Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
  TITLE     Full-length messenger RNA sequences greatly improve genome
            annotation
  JOURNAL   Genome Biol. 3 (6), RESEARCH0029 (2002)
   PUBMED   12093376
REFERENCE   2  (bases 1 to 1273)
  AUTHORS   Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
  TITLE     Full-length cDNA from Arabidopsis thaliana
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1273)
  AUTHORS   Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
            Feldmann,K.
  TITLE     Direct Submission
  JOURNAL   Direct Submission
            Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
            Malibu, CA 90265, USA
COMMENT     This clone sequence is one of 5,000 Ceres full-length cDNAs made
            available to TIGR and Genbank. The following quality assessment of
            this set was done by comparison with known proteins: two percent of
            the clones are estimated to be 5'-truncated; less than one percent
            are 3'-truncated; approximately two percent represent alternative
            splice variants, including unspliced introns and spliced exons; one
            percent may contain premature stop codons; five percent may have
            frame shifts in a coding region. A sequence is considered to be
            5'-truncated if it lacks the translation initiation start (ATG). A
            sequence is considered to be 3'-truncated if it lacks the
            C-terminal end of the encoded protein. Please note that these cDNA
            sequences are derived from the Ws or LAer ecotypes and therefore
            may contain polymorphisms when compared to sequences from Col-0.
            Genset carried out the library production and sequencing of the
            full-length clones. Ceres, Inc. carried out the clustering of the
            5' sequences, selection of clones, and sequence assembly.
FEATURES             Location/Qualifiers
     source          1..1273
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="124067"
     CDS             66..1145
                     /codon_start=1
                     /product="steroid sulfotransferase-like protein"
                     /protein_id="AAM61557.1"
                     /db_xref="GI:21537216"
                     /translation="MATSSMKSIPMAIPSFSMCHKLELLKEGKTRDVPKAEEDEGLSC
                     EFQEMLDSLPKRGWRTRYLYLFQGFWCQOAKEIQAIMSFQKHFQSLENDVVLATIPKS"
ORIGIN
            GTTWLKALTFTLLNRHRFPDVPVASSTNPLFTSNPHDLVFFFEYKLYANGDVPDLSGDLA
            SPRTFATHLPFGSLKEIEKPGVKVVVLCRNPDTFISSWHYTNNIKSESSVSPVLLDQ
            AFDLYCRGVIGFQPFWEHMLGTWRESLSRPPEKVYFFLRYEDLKDDIETNLKRLATFLEL
            PFTEEERKGVVKAIAELCSFENLKCLEVNKNKSIKQNFENRFLFRKGEVSDWNYLS
            PSQVERLSALVDDKLGGSGLTFPLS"
```

US-10-019-931-3 (1-359) x AY084999 (1-1273)

Alignment Scores:

```
Pred. No.:                1
Score:             1.72e-156        Length:           1273
                     1898.00        Matches:           359
Percent Similarity:  100.00%        Conservative:        0
Best Local Similarity: 100.00%      Mismatches:          0
Query Match:         100.00%        Indels:              0
DB:                       15        Gaps:                0
```

```
Qy      1 MetAlaThrSerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHis 20
Db     66 ATGGCGACCTCAAGCATGAAGAGCATTCCAATGGCGATCCCAAGTTTCTCCATGTGTCAC 125

Qy     21 LysLeuGluLeuLeuLysGluGlyLysThrArgAspValProLysAlaGluGluAspGlu 40
Db    126 AAGCTCGAGCTCCTTAAAGAAGGCAAAACTCGCGACGTCCCGAAAGCCGAAGAAGATGAA 185

Qy     41 GlyLeuSerCysGluPheGlnGluMetLeuAspSerLeuProLysArgGlyTrpArg 60
Db    186 GGGCTAAGCTGCGAGTTCCAAGAGATGTTGGATTCTCTTCCTAAGGAGAGGAGGATGGAGA 245

Qy     61 ThrArgTyrLeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGluIleGlnAlaIle 80
Db    246 ACTCGTTACCTTTACCTATTCCAAGGGTTTTGGTGCCAAGGCAAAGAGATTCAAGCCATC 305

Qy     81 MetSerPheGlnLysHisPheGlnSerLeuGluAsnAspValValLeuAlaThrIlePro 100
Db    306 ATGTCTTTCCAAAAACATTTCCAATCCCTTGAAAACGACGTCGTTCTCGCCACATACCT 365

Qy    101 LysSerGlyTyrThrThrTrpLeuHisPheLeuLeuAsnArgHisArgPhe 120
Db    366 AAAATCCGGTACAACCTGGCTAAAAGCTTTAACTTCACCATCCTTAACCGTCACCGGTTT 425

Qy    121 AspProValAlaSerSerThrAsnHisSerProLeuProPheThrSerAsnProHisAspLeuVal 140
Db    426 GATCCGGTTTGCCTCGAGTACCAACCACCCTCTTTCACTTCCAACCCTCATGACCTTGTA 485

Qy    141 ProPhePheGluTyrGlyLysLeuTyrAlaAsnGlyAspValProAspLeuSerGlyLeuAla 160
Db    486 CCTTTCTTCGAGTACGGAAAGCTTTACGCCAACGGAGATGTTCCCGATCTCTCGGGTCTAGCC 545

Qy    161 SerProAlaThrPheAlaThrHisLeuProPheGlySerLeuLysGluIleGluLys 180
Db    546 AGTCCAGCAACGTTCGCAACCCACTTACCGTTCGGTTCTTTAAAGGAAATCGAGAAA 605

Qy    181 ProGlyValLysValValValLeuCysArgAsnProAspThrPheIleSerSerTrp 200
Db    606 CCCGGTGTGAAAGTCGTGTACTTGTGCCGGAACCCGTTTGACACATTCATCTCTTCGTGG 665

Qy    201 HisTyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaPhe 220
Db    666 CATTACACCAACAACATCAAAATCCGAGTCAGTGAGCCCAGTCTTGCTAGACCAAGCTTT 725

Qy    221 AspLeuTyrCysArgGlyValIleGlyPheGlnProPheTrpGluHisMetLeuGlyTyr 240
Db    726 GATCTGTATTGCCGGGGAGTGATCGGATTCCAGCCTTTTTGGGAACACARGTTGGGATAC 785

Qy    241 TrpArgGluSerLeuSerArgProLysValPhePheLeuArgTyrGluAspLeuLys 260
Db    786 TGGAGAGAGAGCTTGAAGACAGAGAAAGTGTTCTTCTTTTAAGGTACGAGGATCTCAAA 845

Qy    261 AspAspIleGluThrAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGlu 280
Db    846 GACGACATCGAGACAAACTTGAAGAGAGGGCTTGCAACTTTCTTAGAGCTTCCTTTCACGGAA 905
```

RESULT 6

```
LOCUS       AY099809                1347 bp    mRNA    linear   PLN 06-MAY-2002
DEFINITION  Arabidopsis thaliana steroid sulfotransferase-like protein
            (At5g07010) mRNA, complete cds.
ACCESSION   AY099809
VERSION     AY099809.1  GI:20466685
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1347)
  AUTHORS   Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,
            Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
            Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
            Kawai,J., Kim,C.J., Liu,S.X., Narusaka,M., Pham,P.K.,
            Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
            Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-APR-2002) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
COMMENT     e-mail for correspondence: arab@sequence.stanford.edu

            RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA ; 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Nguyen,M.,
            Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,
            Bowser,L., Jones,T., Chen,H., Cheuk,R., Chung,M.K.,
            Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
            Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

            Nguyen,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
            equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
            (SSP/Stanford) contributed equally to this work as Pis.
FEATURES             Location/Qualifiers
     source          1..1347
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="RAFL09-95-D03 (R25564)"
                     /ecotype="Columbia"
                     /note="This clone is in pBluescript"
     gene            1..1347
                     /gene="At5g07010"
     CDS             66..1145
                     /gene="At5g07010"
                     /codon_start=1
                     /product="steroid sulfotransferase-like protein"
                     /protein_id="AAM20660.1"
                     /db_xref="GI:20466686"
                     /translation="MATSSMKSIPMAIPSFSMCHKLELLKEGKTIRDVPKAEEDEGLSC
                     EFQEMLDSLPKERGWRTRYLYLFQGPWCQAKEIQAIMSFQKHFQSLBNDVLATIPKS
                     GTTWLKALIFTILNRHFRPDVASSTNHFLFTSNPHDLVPFFEYKLYANGDVPDLSGIA
                     SPRTFATHLPFGSLKETIEKPGVKVVVLCRNPFDTFISSWHYINNIKSESVSPVLLDQ
                     AFDLYCRGVIGFGPFWEHNLGYWRESLRRAPEKVFFLRYEDLKDDIEINLATFLEL
                     PFTEEEERGVVKAIAELCSFENLKKLEVNKSNKSIKNFENRFLFRKGEVSDWNYLS
                     PSQVERLSALVDDKLGGSGLTFRLS"
ORIGIN
```

```
US-10-019-931-3 (1-359) x AY099809 (1-1347)

Alignment Scores:
Pred. No.:
                       1.84e-156    Length:          1347
Score:                 1898.00      Matches:         359
Percent Similarity:    100.00%      Conservative:    0
Best Local Similarity: 100.00%      Mismatches:      0
Query Match:           100.00%      Indels:          0
                       15           Gaps:            0
DB:
```

```
US-10-019-931-3 (1-359) x AY099809 (1-1347)

Qy      1 MetAlaThrSerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHis 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     66 ATGGCTACCTCAAGCATGAAGAGCATTCCAATGGCCATCCCAAGTTTCTCCATGTGTCAC 125

Qy     21 LysLeuGluLeuLeuLysGluGlyLysThrArgAspValProLysAlaGluGluAspGlu 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    126 AAGCTCGAGCTCCTTAAAGAAGGGAAAACTCGCGACGTCCCCGAAAGCCGAAGAAGATGAA 185

Qy     41 GlyLeuSerCysGluPheGlnGluMetLeuAspSerLeuProLysGluArgGlyTrpArg 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    186 GGGCTAAGCTGCGAGTTCCAAGAGATGTTGGATTCTCTTCCTAAGGAGAGAGGATGGAGA 245

Qy     61 ThrArgTyrLeuTyrLeuPheGlnGlyProTrpCysGlnAlaLysGluIleGlnAlaIle 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    246 ACTCGTTACCTTTACCTATTCCAAGGGCCTTTGGTGCCAAGCAAAGAGATTCAAGCCATC 305

Qy     81 MetSerPheGlnLysHisPheGlnSerLeuGluAsnAspValLeuAlaThrIlePro 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    306 ATGTCTTTCCAAAAACATTTCCAAATCCCTCGAAAACCCACGTCGTTCCGCCACCATACCT 365

Qy    101 LysSerGlyThrThrTrpLeuLysAlaLeuThrPheThrIleLeuAsnArgHisArgPhe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    366 AAATCCGGTACAACCTGGCTAAAAGCTTTAACTTTCACAATCCTTAACCGTCACCGGTTT 425

Qy    121 AspProValAlaSerSerThrAsnHisPheLeuPheThrSerAsnProHisAspLeuVal 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    426 GATCCGGTTGCCTCGAGTACCAACCACCCTCTTTTCACTTCCAACCCTCATGACCTTGTA 485

Qy    141 ProPhePheGluTyrLysLeuTyrAlaAsnGlyAspValProAspLeuSerGlyLeuAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    486 CCTTTCTTCGAGTACAAGCTTTACGCCAACGGAGATGTTCCGGACCTCTCGGGTCTAGCC 545

Qy    161 SerProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluLys 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    546 AGTCCAAGAACGTTCGCAACCCACTTACCGTTCGGTTCCCTTAAAGGAAACGATCGAGAAG 605

Qy    181 ProGlyValLysValValValLeuCysArgAsnProPheAspThrPheIleSerSerTrp 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    606 CCCGGTGTGAAGGTCGTTGTACTTGTCCGGAACCCGTTTGACACATTCATCTCTTCGTGG 665

Qy    201 HisTyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaPhe 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    666 CATTACACACCAACAACATCAAATCCGAGTCAGTGAGCCCAGTCCTTGCTTGACCAAGCTTT 725

Qy    221 AspLeuTyrCysArgGlyValIleGlyPheGlyProPheTrpGluHisMetLeuGlyTyr 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    726 GATCTCTATTGCCGCGGAGTGATCGGGTTTGGCCCGTTTTGGGAACACATGTTGGGATAC 785

Qy    241 TrpArgGluSerLeuArgArgAlaProGluLysValPhePheLeuArgTyrGluAspLeuLys 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    786 TGGAGAGAGAGCTTGAGGAGAGACCAGAGAAAGTCTTCTTTTTAAGGTACGAGGATCTCAAA 845
```

```
Qy    281 GluGluGluArgLysGlyValValAlaValAlaIleAlaGluLeuCysSerPheGluAsnLeu 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    906 GAAGAGGAACCGAAAGGAGAGTTGTGAAGCGTATCGCCGAGCTGTGTAGCTTCGAGAATCTG 965

Qy    301 LysGluLeuGluValAsnLysSerAsnLysIleLysAsnPheGluAsnArgPheLeu 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    966 AAGGAGTTGGAGGTGAACAAGTCAAACAAGTCGATCAAGAAACTTTGAGAATCTTG 1025

Qy    321 PheArgGlyGluValSerAspAspTrpValAsnTyrLeuSerProSerGlnValGluArg 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1026 TTTCGGAAAGGAGAAGTGAGTGATGGGTTAACTATTTGTCACCTTCACAAGTGGAAGA 1085

Qy    341 LeuSerAlaLeuValAlaAspAspLysLeuGlyGlySerGlyLeuThrPheArgLeuSer 359
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1086 TTGTCAGCCTTAGTGGATGACAAAGTTAGGTGGATCTGGTCTCACTTTCAGGTTGAGC 1142
```

```
Qy    261 AspAspIleGluThrAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGlu 280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    846 GACGACATCGAGACCAACTTGAAGAGGCTTGCAACTTTCTTAGAGCTTCCTTTCACCGAA 905

Qy    281 GluGluGluArgGlyGlyValValLysAlaIleAlaGluLeuCysSerPheGluAsnLeu 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    906 GAAGAGGAACGAGGAAAGGGAGAGTTGTGAAGGCTATCGCCGAGCTGTGTAGCTTCGAGAATCTG 965

Qy    301 LysLysLeuGluValAlaAsnLysSerAsnLysSerAsnLeuIleLysAsnMetArgPheLeu 320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    966 AAGAAGTTGGAGGTGGCAAACAAGTCAAACAAGTCGATCAAGAACATTATGAGAATCGAATTCTTG 1025

Qy    321 PheArgLysGlyGluValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArg 340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1026 TTTCGGAAAGGAGAAGTGAGTGACTGGGTTAACTATTTGTCACCTTCACAAGTGGAAAGA 1085

Qy    341 LeuSerAlaLeuValAlaSpAspLysGlyGlySerGlyLeuThrPheArgLeuSer 359
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   1086 TTGTCAGCCTTAGTGGATGACAACAAGTTAGGTGGATCCGGTCTCACTTTCAGGTTGAGC 1142
```

```
RESULT 7
LOCUS        AB010697              86380 bp    DNA    linear   PLN 14-FEB-2004
DEFINITION   Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MOJ9.
ACCESSION    AB010697
VERSION      AB010697.1  GI:2828182
KEYWORDS     .
SOURCE       Arabidopsis thaliana (thale cress)
  ORGANISM   Arabidopsis thaliana
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
             rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1
  AUTHORS    Kaneko,T., Kotani,H., Nakamura,Y., Sato,S., Asamizu,E., Miyajima,N.
             and Tabata,S.
  TITLE      Structural analysis of Arabidopsis thaliana chromosome 5. V.
             Sequence features of the regions of 1,381,565 bp covered by twenty
             one physically assigned P1 and TAC clones
  JOURNAL    DNA Res. 5 (2), 131-145 (1998)
   PUBMED    9679902
REFERENCE    2  (bases 1 to 86380)
  AUTHORS    Nakamura,Y.
  TITLE      Direct Submission
  JOURNAL    Submitted (28-JAN-1998) Yasukazu Nakamura, Kazusa DNA Research
             Institute, Department of Plant Gene Research; 1532-3, Yana,
             Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
             Tel:81-438-52-3935, Fax:81-438-52-3934)
COMMENT      Address for correspondence: kaos@kazusa.or.jp
             For the latest information on annotation of this clone, please see
             http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MOJ9
             Genes with similarity to proteins in the databases are described in
             'product', or 'note' qualifiers. Genes that have no significant
             protein similarity are described as 'unknown protein'.
             The software programs used to predict genes include: Grail
             (Informatics Group, Oak Ridge National Laboratory,
             http://compbio.ornl.gov/Grail-1.3/),
             GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
             NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
             Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
             SplicePredictor (Volker Brendel, Stanford University,
             http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi).
             Genes encoding tRNAs are predicted by tRNAscan-SE
             (Sean Eddy, Washington University School of Medicine, St. Louis,
             http://genome.wustl.edu/eddy/tRNAscan-SE/).
             This sequence may not be the entire insert of this clone. It may be
             shorter because we remove overlaps between neighboring submissions.
             The 5' clone is MPH15 and the 3' clone is T28J14.
FEATURES             Location/Qualifiers
     source          1..86380
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
                     /chromosome="5"
```

```
     CDS             complement(6176..8560)
                     /note="gene_id:MOJ9.2"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="anthranilate phosphoribosyltransferase-like
                     protein"
                     /protein_id="BAB11143.1"
                     /db_xref="GI:9759541"
                     /translation="MAAKDGAKSQEDYKLKDMKPELGERWPHGQORGGTGWIGGSERAA
                     STVDLVEQMFYLYRVKVKAKDLPPNVPVTSNCDPYVEVKIGNYKGKTKHFEKRTNPEWN
                     QVPAFSKDKVQSSTVEVFVRDKEMVRDEYIGKVVFDKREVFPIVWPPDSPLAPQWYRL
                     EDRRGBESKRGEWVVAWNLGTQADEAFPDAWHSDASSVGGEGVGQSVRSKVVSEKLWY
                     LRVNVIEAQDVEFPSDRSQPPQAFVKVQVGNQILKTKLCPNKTTNPMWNEDLVFAAEP
                     QLAVRFTCLSLAHMIYLYGHPLLPKMHVLHPFTVNQLDSIRYQAMSIVAARLSRAEPP
                     LRKENVEYMLDVDSHMWSMRRSKANFFRIVSVFAGLIAMSKWLGDVCYWKNPLTTILF
                     HVLFFILICYPEIILPTTFLYMFLIGLWMNFRFRPHPAHMDTKYSWAARAASPDELDEE
                     FDIPPTSKGQDVVKMRYDRLRSVAGRIQMVVGDIATQGERFQALLSWRDPRATCLFVI
                     LSPMKTKDGKAITDPYCVAKYGQKWVRTRTIIDDSSSPKWNEQYTWEVIDPCTVILGV
                     FDNCHLGGSEKSNSGAKVDSRIGKVRIRLSTLEADRIYTHSYPLLVLQTKGLKMGEV
                     join(11149..11889,11959..12410)"
                     /gene_id="gb|AAF69827.1"
```

```
     CDS             join(11149..11889,11959..12410)
                     /note="gene_id:MOJ9.3"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="polygalacturonase inhibiting protein 1; PGIP1"
                     /protein_id="BAB11144.1"
                     /db_xref="GI:9759542"
                     /translation="MDKTRTLCLLFLFTFITTCLSKDLCNQNDKNTLLKIKKSLNNPY
                     HLASWDPQTDCCSWYCLECGDATVNHRVTALTITGSGQISGQISGQIPEVGDLPYLETLVFR
                     KLSNITGTIQPTIAKLKNLRMLRLSWTNLTGPIPDFIGQLKNLEFLELSFNDLSGSIP
                     SSLSTLPKILALELSRNKLTGSIPESFGSFPGTVPDLRLSHNQLSGPIPKSLGNIDFN
                     RIDLSRNKLQGDASMLFGSNKTTWSIDLSRNMFQFDISRVDIFRKTIGILDLNHNGITG
                     NIPVQWTEAPLQFFNVSYNKLCGHIPTGCKLQTFDSYSFFHNKCLCGAPLEICK"
```

```
     CDS             join(12917..13457,13541..13992)
                     /note="gene_id:MOJ9.4"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="polygalacturonase inhibiting protein"
                     /protein_id="BAB11145.1"
                     /db_xref="GI:9759543"
                     /translation="MDKTWTLFLLFLLSTLLLTTSLAKDLCHKDKDKTTLLKIKKSLNNPY
                     HLASWDPKYDCCSWYCLECGDATVNHRVTSLLIIODGEISGQIPEVGDLPYLTSLIFR
                     KLTNLTGHIQPTIAKLKNLRLSWTNLTGPIPESFGTFPSGKVPSLLFLSHNQLSGPIPK
                     SSLSLIRKLEYLELSRNKLTGPIPESFGTFPSGKVPSLFLSHNQLSGPTPKSLGNPDFY
                     RIDLSRNKLQGDASILFGAKKTTWIVDISRNMFQFDLSKVKLIAKTLANNLDMNNGITG
                     SIPAEWSKAYFQLLNVSYNRLCGRIPKGEYIORFDSYSFFHNKCLCGAPLPSCK"
```

```
     gene_id="unnamed protein product; gb|AAF63169.1
             14719..14919
                     /note="gene_id:MOJ9.5
             14719..14919
                     similar to unknown protein"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="BAB11146.1"
```

```
     CDS             join(1077..1154,1263..1322,1410..1487,1579..1624,
                     1754..2025,3170..3424,5041..5193)
                     /codon_start=1
                     /note="gene_id:MOJ9.1"
                     /evidence=not_experimental
                     /product="bZIP transcription factor-like protein"
                     /protein_id="BAB11142.1"
                     /db_xref="GI:9759540"
                     /translation="MLWQKEGNRKGLASSDHDIPKSSDPKTLRRLAQNREAARKSRLR
                     KKAYVQQLESCRIKLTQLEOEIQRARSQGVFGGSLIGDQQOEAAVFDMEYARWLEE
                     QORLINELRVATOEHLSENELRMFVDTCLAHYDHLINLKAMVARKTDVFHLISGAWKTP
                     AERCFLWMGGFRPSEIIKVIVNQIEPLTEQQIVGICLQOSTQEAEEALSOGLEALNQ
                     SLSDSIVSDSLPPASAPLPPHLSNFMSHMSLALNKLSALEGFVLQADNIRHQTIHRLN
                     QLILTRQEARCLLAVAEYFHRLQALSSLWLARPRODG"
```

CDS
/db_xref="GI:9759544"
/translation="MVSVPDVNDGQVIEITVQSLSENVASLKEKVANIKLIRGKVAGIL
KDDDKSLAHYNVRAGDILFLSW"
/codon_start=1
/evidence=not_experimental
/product="cytochrome P450"
/protein_id="BAB11147.1"
/db_xref="GI:9759545"
/note="gene_id:MOJ9.6"
complement(join(15136..15762,16005..16901))

CDS
/translation="MVDLQYFSVIILVCLGITTVLIQAITNRLRDRLPLPPSPTALPII
GHIHLLGPIAHQALHKLSRYGPLMYLFIGSIPNLIVSSAEMANEILKSNELNPINRP
TMQNVDYLITYGSADFFSAPYGLHWKFMRICIMCWELFSSRALDSFVSRSEELKLLIR
VLKKAEAESVNLGEQLKELTSNIITRMWFRKMQSDSDGGEKSEEVIKMVELNBLAG
FFNVSETFWFLKRLDLQGLKKRLKNARDKYDVITERIMEHESSKKNATGERNMLDVL
LDIYEDKNAEMKLTRENIKAFIMNIYGGGTDTSAITVEWALALINHPEIMKKAQQEI
EQVGNKRVVEESDLCNLSYTQAVVKETMRLHPGGPIFVRESDBECAVAGFRIPAKTR
VIVNWAIGRDSNQWEDPLEFRPRPFEGGSEWKVMSEKMMSFGAGRRSCPGEKMVFRFV
PILLAAIIQCFELKVKGSVDMEDGTGSSLPRATPLVCVPVAKEATQSFSLLIENVNF"
/codon_start=1
/evidence=not_experimental
/note="gene_id:MOJ9.7"
complement(join(17414..18091,18167..19075))

CDS
/db_xref="GI:9759546"
/translation="MFTLNIDMNYTSCSYLFFILVTIFILLHRLFSSSSRRGLPPGPRG
LPVLGHMHLLRSSLPRSLOALAHTYQLMTIRIGSLRVLVVSSDSDTAKLILKTHDPDF
ASKRVFGPRQFNVVKGSEFNAPYGSYWFMKKLCMTKLFAGYLQLDRFVDIREEETLA
LSTLVERSRNGEACDLGLEFTALITKILSKWVMGKRCRQNSNIPKEIRKIVSDIMAC
ATRRGFMELFGPLRDLDLFGNGNGKKLRSIIWRYDELVEKLIKEYENDKSNEEEERDKDI
VDILLDTYNDPKAELRLTMNQIKFFILELFMASLDTTSAALQWTMTELINHPDIFAKI
RDEIKSVVGTTNRLIKESDLQKLPYLQAAIKETLRLHPVGPLLRRESNTDMKINGYDV
KSGTKIFINAYGIMRDPTTYKDDPDKFMPERFLVVEQDTERKMGYYQQWMLELKGQDVN
YLARGSGRRGCLGASHASIVLSITIGSLVQCFNWTVKGDEDKIKLPTGFSASGTAG
GSSLMCSPELCFDPFGYQTK"
/codon_start=1
/evidence=not_experimental
/note="gb|AAB91418.1|"
join(19574..19666,19980..20078,20155..20256,20373..20432,
20524..20630,20702..20784,21103..21199,21291..21501)

CDS
gene_id:MOJ9.8"
/codon_start=1
/evidence=not_experimental
/product="DnaJ homologue"
/db_xref="GI:9759547"
/protein_id="BAB11149.1"
/translation="MGRKKKSRASTTEDEIEMDNAGPSSSETSLYEVLGVERATSQE
IRKAYHKLALKLHPDKNQDDKEAKDKFQQLKQLKVTSILGDEEKRAVYDQTGSIDDADIP
GDAPENLRDFPRDMYKKVNEADIEEFEANYRGSSESEKDLLELFNKFKGKQMNRLFCSM
LCSDPKLDSHRFKDMLDEAIAAGEVKSKSAYEKWANKLSETKPTSPLRKRKKKKSAA
KDSETDLCLMIAKRQBERRGKVDSMFSSLISRYCGDAEAEPTEEEFEAAQRRIESKRK
PSKKSRGK"
/codon_start=1
/evidence=not_experimental
/note="gene_id:MOJ9.9
similar to unknown protein"
gene_id:MOJ9.10
/note="unnamed protein product; gb|AAD39647.1
21834..22895
join(24115..24544,24743..25145,25239..25425,25504..25713,
25795..25968,26052..26216,26328..26441,26535..26825)

CDS
/translation="MGCCSSDCFYVFILSIALAPMAISTTLRSPPDSEPTIPIAFSSS
SPSLSLNASNTLRQSNFKAIATLLHISPEIFLSSSPNTTLFAIEDASFNTSSLHPLF
LKQLLHYHTPLMLSMDDLLKKPQGTCLPTLLHHKSVQISTVNQESRTAENVHVRITH
PDMFLGDSLVIHGVIGFFSPLQPHSDHLIHTPLCQSDTTNKTSNNEEVPVSIDMTRIV
QLLSSNGFVPFAIGLHSVLANRIVANDHHNHKNLLTGVTILATPNLVSLSASPFLYEVVR
HHILVQRLFYKDFASMSDKATVKTLDPYQDLITIRRNVNNSSGGDFMISGVEIVDPDMF
SSNFVIHGISHTLEIPHV"
/codon_start=1
/evidence=not_experimental

US-10-019-931-3 (1-359) x AB010697 (1-86380)

Alignment Scores:

Pred. No.:	1	
Score:	1898.00	Length: 86380
Percent Similarity:	100.00%	Matches: 359
Best Local Similarity:	100.00%	Conservative: 0
Query Match:	100.00%	Mismatches: 0
DB:	15	Indels: 0
		Gaps: 0

/protein_id="BAB11151.1"
/db_xref="GI:9759549"
/translation="MKRFERPRTSIHSQPRTSTSSGKKDQKPRSSSSKGSGVKPLNAA

Qy   1 MetAlaThrSerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHis 20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 55015 ATGGCTACCTCAAGCATGAAGAGCATTCCAATGGCCATCCCAAGTTTCTCCATGTGTCAC 54956

Qy   21 LysGluLeuLeuLeuLysGluGlyLysThrArgAspValProLysGlyLysGluGluGlu 40
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54955 AAGCTCGAGCTCCTTAAAGAAGGGCAAAACTCGCAGACGTCCGAAAGCGAAGAAGAGAA 54896

Qy   41 GlyLeuSerCysGluGluMetLeuAspSerLeuProLysGluGlyLysTyrTrpArg 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54895 GGGCTAAGCTCCGAGTTTCCAAGCATGAAGAGCATTCCAATGGCCATCCCAAGTTTCTCCAT 54836

Qy   61 ThrArgTyrLeuTyrIleLeuPheGlnGlyPheTrpCysValArgAlaAsnGlyAla 80
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54835 ACTCGTTACCTTTACCTATTCCAAGGGTTTTGGTGTCAAGAGCAAGAGAATTCAAGGCATC 54776

Qy   81 MetSerPheGlnLysHisPheGlnSerLeuGluAsnAspLysLeuAlaThrIlePro 100
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54775 ATGTCTTTCCAAAAACATTTCCAATCCCTCGAAAACGACAAGCTCGCCACAATACCT 54716

Qy   101 LysSerGlyThrThrTrpLeuLysAlaLeuThrPheThrIleLeuLeuAsnArgHisArgPhe 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54715 AAAATCCGGTACAACCTGGCTAAAAGCTTTAACTTTCACCATCCTTAACCGGTTT 54656

Qy   121 AspProValAlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAspLeuVal 140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54655 GATCCGGTTGCCTCGAGTACAAACCACCCTCTTTTCACTTCAAACCCTCATGACCTTGTA 54596

Qy   141 ProPhePheGluTyrLysLeuTyrIleAlaAsnGlyAspValProAspLeuSerGlyLeuAla 160
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54595 CCTTTCTTCGAGTACAAGCTTTACCGAGATGTTCCCGGATCTCGGGTTGCAGCC 54536

Qy   161 SerProAsnArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluLys 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54535 AGTCCAAGAACGTTCCCAACCCACTTACCGTTCGGTTCTTAAAGGAAACGATCGAGAAA 54476

Qy   181 ProGlyValLysValThrTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrp 200
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54475 CCCGGCGTCGTAGGTCGTACTTACTTGTGCCGGAACCCGTTGACACATTCATCTTCGTGG 54416

Qy   201 HisTyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaPhe 220
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54415 CATTACACCAACAACATCAAATCCGAGTCAGTGAGCCCGGTCCTGCTAGACCAAGCTTTT 54356

Qy   221 AspLeuTyrCysArgGlyValIleGlyPheGlyTrpGluHisMetLeuGlyTyr 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54355 GATCTGTATTGCCGGGGAGTGATCGGGTTTGGGTGGGAACACATGTTGGGATAC 54296

Qy   241 TrpArgGluSerLeuLysArgProGluLysValPhePheLeuArgTyrGluAspLeuLys 260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54295 TGGAGAGAGAGCTTGAAGAGACCAGAGAAAGTCTTCTTTTTAAGGTACGAGGATCTCAAA 54236

Qy   261 AspAspIleGluThrPheAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGlu 280
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54235 GACGACATCGAGACCAACTTGAAGAGGCTTGCAACTTTCTTGGAGCTTCCTTTCACGGAA 54176

Qy   281 GluGluGluArgLysGlyValValLysAlaIleAlaGluLeuCysSerPheGluAsnLeu 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54175 GAAGAGGAACGAAAGGGAGTGGTTGTGAAGGCTATACGCGAGCTGTGTAGCTTCGAGAATCTG 54116

```
Qy   301 LysLysLeuGluGluValAlaAsnLysSerAsnLysSerIleGlnAsnPheGluAsnArgPheLeu 320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54115 AAGAAGTTGGAGGAGGTGGCAAACAAGTCAAACAAGTCGATCAAGAACAAGTTTGAGAATCGATTCCTTG 54056

Qy   321 PheArgLysGlyGluValSerAspTrpAsnTyrLeuSerProSerGlnValGluArg 340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 54055 TTTCGGAAAGGAGAAGTGAGTGATTGGAATTATTTGTCACCTTCACAAGTGGAAAGA 53996

Qy   341 LeuSerAlaLeuValAspAspLysLeuGlyGlySerGlyLeuThrPheArgLeuSer 359
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 53995 TTGTCAGCCTTAGTGGATGACAAGTTAGGTGGATCTGGTCTCACTTTCAGGTTGAGC 53939
```

RESULT 8

```
LOCUS       AY054219                1270 bp    mRNA    linear   PLN 30-SEP-2001
DEFINITION  Arabidopsis thaliana AT5g07010/MOJ9_18 mRNA, complete cds.
ACCESSION   AY054219
VERSION     AY054219.1  GI:15809902
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1270)
  AUTHORS   Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Koesema,E., Meyers,M.C.,
            Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D.,
            Hayashizaki,Y., Jiang,P.X., Jones,T., Kamiya,A.,
            Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X.,
            Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
            Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A.,
            Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
            Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
  TITLE     Arabidopsis cDNA clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1270)
  AUTHORS   Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Koesema,E., Meyers,M.C.,
            Banh,J., Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D.,
            Hayashizaki,Y., Ishida,J., Jiang,P.X., Jones,T., Kamiya,A.,
            Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X.,
            Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
            Pham,P.K., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A.,
            Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
            Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-2001) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA) : Seki,M.,
            Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T.,
            Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.
            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Shinn,P., Chen,H.,
            Cheuk,R., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J., Bowser,L.,
            Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
            Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
            Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
            Davis,R.W., Theologis,A. and Ecker,J.R.
            Shinn,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1..1270
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="RAFL09-22-M04 (R13921)"
                     /ecotype="Columbia"
     5'UTR           1..67
     CDS             68..1147
                     /note="steroid sulfotransferase-like protein"
                     /codon_start=1
                     /product="AT5g07010/MOJ9_18"
                     /protein_id="AAL06879.1"
                     /db_xref="GI:15809903"
                     /translation="MATSSMKSIPMAIPSFSMCHKLELLKEGKTRDVPKAEEDEGISC
                     EFQEMLDSLPKERGWRTRYLYLFQGFWCQAKEIQAIMSFQKHFQSLENDVLAIIPKS
                     GTTWLKALTFTLNRHRFDPVASSTNHHLFTSNPDLVFPFEYKLYANGDVPDLSGLA
                     SPRTFATHLPFGSLKETIEKPGVAVVYLCRNPFDTFISSWHYTNNIKSESVSPVLLDQ
                     AFDLYCRGVIGFAPFWEHMLGYWRESLKRPEKVFFLRYEDLKDDIETNLKRLATFIEL
                     PFTEEEERKGVVKAIAELCSFENLKLEVNKSNKSIQNFENRFLFRKGEVSDWNYLS
                     PSQVERLSALVDDKLGGSGLTFRLS"
     3'UTR           1148..1270
ORIGIN
```

US-10-019-931-3 (1-359) x AY054219 (1-1270)

```
Alignment Scores:
Pred. No.:
Score:                    1892.00     Length:          1270
Percent Similarity:         99.72%    Matches:          358
Best Local Similarity:      99.72%    Conservative:       0
Query Match:                99.68%    Mismatches:         0
DB:                                   Indels:             1
                                      Gaps:               0
```

```
Qy    1 MetAlaThrSerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHis 20
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   68 ATGGCTACCTCAAGCATGAAGAGCATCCCAATGGCGATCCCAAGTTTCTCCATGTGTCAC 127

Qy   21 LysLeuGluLeuLeuLysGluGlyLysThrArgAspValProLysAlaGluGluAspGlu 40
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  128 AAGCTCGAGCTCCTTAAAGAAGGCAAAAACTCCGGACGTTCCGAAAGCCGAAGAAGATGAA 187

Qy   41 GlyIleSerCysGluPheGlnGluMetLeuAspSerLeuProLysGluArgGlyTrpArg 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  188 GGGCTAAGCTGCGAGTTCCAAGAGATGTTGGATTCTCTTCCTAAAGGAGAGAGGGGATGGAGA 247

Qy   61 ThrArgTyrLeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGluIleGlnAlaIle 80
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  248 ACTCGTTACCTTTACCTATTCCAAGGGTTTTGGTGCCAAGCAAAGATTCAAGCCATC 307

Qy   81 MetSerPheGlnLysHisPheGlnSerLeuGluAsnAspValLeuAlaThrIlePro 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  308 ATGTCTTTCCAAAAACATTTCCAATCCCTCGAAAACGACGTGCTTGCCACCATACCT 367

Qy  101 LysSerGlyThrThrTrpLeuLysAlaLeuThrPheThrLeuAsnArgHisArgPhe 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  368 AAATCCGGTACAACCTGGCTAAAAGCTTTAACTTTCACCCTAAACCGTCACCGGTTT 427

Qy  121 AspProValAlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAspLeuVal 140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db  428 GATCCGGTTGCCTCGAGTACCAACCACCCTCTTTTCACTTCACCAAACCCTCATGACCTTGTA 487

Qy  141 ProPhePheGluTyrLysLeuTyrAlaAsnGlyAspValProAspLeuSerGlyLeuAla 160
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  488 CCTTTCTTCGAGTACAAGCTTTACGCCAACGGAGATGTTCCGGATCTCTCGGGTCTAGCC 547

Qy  161 SerProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluLys 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  548 AGTCCAAGAACGTTCGCCACACACTTACCGTTCGGTTCCTTAAAGGAAACGATAGAGAAG 607

Qy  181 ProGlyValValValTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrp 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  608 CCCGGTGTGGAGGTCGTGTACTTGTGCCGGAACCCGTTTGACACATTCATCTCTTCGTGG 667

Qy  201 HisTyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaPhe 220
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  668 CATTACACCAACAACATCAAAAATCCGAGTGAGCCCAGTACTTGCTAGACCAAGCTTTT 727
```

RESULT 9

```
LOCUS       AK117202              1236 bp    mRNA    linear   PLN 14-FEB-2004
DEFINITION  Arabidopsis thaliana At5g07000 mRNA for putative steroid
            sulfotransferase, complete cds, clone: RAFL16-75-F02.
ACCESSION   AK117202
VERSION     AK117202.1  GI:26449503
KEYWORDS    FLI cDNA; CAP trapper.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Seki,M., Iida,K., Satou,M., Sakurai,T., Akiyama,K., Ishida,J.,
            Nakajima,M., Enju,A., Kamiya,A., Narusaka,M., Carninci,P.,
            Kawai,J., Hayashizaki,Y. and Shinozaki,K.
  TITLE     Arabidopsis thaliana full-length cDNA
  JOURNAL   Published Only in Database (2002)
REFERENCE   2  (bases 1 to 1236)
  AUTHORS   Seki,M., Iida,K., Satou,M., Sakurai,T., Akiyama,K., Ishida,J.,
            Nakajima,M., Enju,A., Kamiya,A., Narusaka,M., Carninci,P.,
            Kawai,J., Hayashizaki,Y. and Shinozaki,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences
            Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa
            230-0045, Japan (E-mail:mseki@gsc.riken.go.jp,
            URL:http://pfgweb.gsc.riken.go.jp, Tel:81-45-503-9625,
            Fax:81-45-503-9586)
COMMENT     An Arabidopsis full-length cDNA library was constructed essentially
            as reported previously (Seki et al. (1998) Plant J. 15:707-720;
            Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
            and XhoI was ligated to modified Lambda FLC-1-E vector (Carninci et
            al. (2001) Genomics 77:79-90) digested with BamHI and SalI.
            This clone is in a modified pBluescript vector.
            Please visit our web site (http://pfgweb.gsc.riken.go.jp/) for
            further details.
FEATURES             Location/Qualifiers
     source          1..1236
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="RAFL16-75-F02"
     gene            1..1236
                     /gene="At5g07000/MOJ9_17"
                     /note="common name: thale cress"
                     /ecotype="Columbia"
     CDS             66..1109
                     /gene="At5g07000/MOJ9_17"
                     /codon_start=1
                     /product="putative steroid sulfotransferase"
                     /protein_id="BAC41878.1"
                     /db_xref="GI:26449504"
                     /translation="MAIPSFSMCHKPELLKEGKSSEGQSEEGLSYEFQEMLDSLPKERG
                     RRNRYLYFGGFRCQAKEIQAITSFQKHFQSLPDDVVLATIPKSGTTWLKALTFILT
                     RHRFDPVSSSSDHPLLTSNPHDLVPFEYKLYANGNVPDLSGLASPRTPATHVPFGA
                     LKDSVENPSVKVVVILCRNPFDTIYSMWHYINNTISESVSAVLLDEAFDLYCRGLLIGF
                     GPFWEHMLGYWRRSSLKRPEKVLFLKYEDLKEDIETNLKLLASFLGLPFTEEEEQKGVV
                     KAIADLCSFENLKKLEVNKSSKLIQNYENRFLRKGEVSDLVNNLSPSQVERLSALVD
                     DKLAGSGLIFRLS"
ORIGIN
```

US-10-019-931-3 (1-359) x AK117202 (1-1236)

```
Alignment Scores:
Pred. No.:                                        Length:            1236
Score:                    7.36e-129               Matches:            305
                          1582.00
Percent Similarity:       92.46%                  Conservative:        26
Percent Local Similarity: 85.20%                  Mismatches:          21
Query Match:              83.35%                   Indels:               6
DB:                       15                       Gaps:                 3
```

```
Qy    4  SerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHisLysLeuGlu   23
            ::|||::|:|||:|:||||||||:
Db   45  TCAAGCATAAAAGAGCAHTTCAAATGGCGATCCCAAGTTTCTCCAAGTGTCAGAAGCCCGAG  104

Qy   24  LeuLeuLysGluGlyLysThrArgAspValProValAlaGluGluAspGluGlyLeuSer   43
            |||||||||||||||:::             :::||||||:::
Db  105  CTCCTTAAGGAAGGCAAAAGC----------GAAGGCCAAGAAGAGAAGGGCTAAGC  152

Qy   44  CysGluPheGlnGluMetLeuAspSerLeuProLysGluArgGlyTrpArgThrArgTyr   63
            |||||||||||:|||||||||||:|||||||
Db  153  TACGAGTTCCAAGAGATGTTGGACTCTCTACCTAAGGAGAGAGAGGGACGGGAAATCGTTAC  212

Qy   64  LeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGluIleGlnAlaIleLeuMetSerPhe   83
            |||||||||||||||||||||||||||||:
Db  213  CTTTACTTATTCCAAGGGTTTTGGTGCCAAGCTAAGGAGATTCAAGCTATCACGTCTTTC  272

Qy   84  GlnLysHisPheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLysSerGly  103
            |||||||||||||||:::||||||||||||||||||
Db  273  CAAAAACATTTCAGTCCCTTCCAAGACGACGGTTGTCCTCGCCACCATACCTAAATCTGGC  332

Qy  104  ThrThrThrTrpLeuLysAlaLeuThrPheIleLeuThrArgHisArgPheAspProVal  123
            |||||||||||||||||||||||:::|||||||||||
Db  333  ACAACCTGGTTAAAAGCTTTAACTTTCACCATCCTTACCCGTCATCGGTTTGATCGGGTT  392

Qy  124  ---AlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAspLeuValProPhe  142
               :::|||||||||||||||||||||||:
Db  393  TCCTCATCAAGTTCCCACACCCTCTTCTCACATCCAACCCTCACCGTTACCCTTC  452

Qy  143  PheGluTyrLysLeuTyrAlaAsnGlyAsnValProAspLeuSerGlyLeuAlaSerPro  162
            |||||||||||||||:::||||||||||||||
Db  453  TTCGAGTACAAGCTTTACGCCAACGGAAAGTGTTCCCGATCTCTCGGGTTCTAGCCAGTCCA  512

Qy  163  ArgThrProPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluLysProGly  182
            |||||||||||||||:::||||||||||||||
Db  513  AGAACATTCCGCAACCCCACGTACCGTTCGGTCGCCCTTAAGGATTCCGTCGAGAATCCCAGT  572

Qy  183  ValLysValValTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrpHisTyr  202
            |||||||||||||||||||||||||||||||
Db  573  GTGAAGGTTGTGTACCTGTGCCGGAACCCGTTTGACACATTCATCTCCAGTTGGCATTAC  632

Qy  203  ThrAsnAsnIleLysSerGluSerValSerProValIleLeuLeuAspGlnAlaPheAspLeu  222
            ||||||||||||||||:::|||||||||||:
Db  633  ATCAACAACATCACTTCCGAGTCAGTGAGCGCAGGTCTTGCTGACGAAGCTTTGGATCTA  692

Qy  221  AspLeuTyrCysArgGlyValIleGlyPheGlyProPheTrpGluHisMetLeuGlyTyr  240
            |||||||||||||||||||||||||||||||
Db  728  GATCTGTATTGCCGCGGAGTGATCGGGTTTGGGCCGTTTTGGGAACACATGTTGGGATAC  787

Qy  241  TrpArgSerSerLeuLysArgProGluLysValLeuPheLeuLysTyrGluAspLeuLys  260
            |||||||||||||||||||||||||||||||
Db  788  TGGAGGAGCAGCCTTGAAGAGACCAGAGAAAGTCTTCTTTTTAAGGTACGAGGATCTCAAA  847

Qy  261  AspAspIleGluThrAsnLeuLysLeuLeuAlaThrPheLeuGlyLeuProPheThrGlu  280
            |||||||||||||||||||||||||||||||
Db  848  GACGACATCGAGACCAACTTGAAGAGGCTTGCAACCTTTCTTGGTTTCTTTCACCGGA  907

Qy  281  GluGluGluGluGlnLysGlyValValLysAlaIleAlaGluLeuCysSerPheGluAsnLeu  300
            |||||||||||||||||||||||||||||||
Db  908  GAAGAGGAACGAAAAGGAGGAGTTGTGAAGGCTATCGCGGAGCTGTGTAGCTTCGAGAATCTG  967

Qy  301  LysLysLeuGluValAsnLysSerSerIleLysTyrGluAsnArgPheLeu  320
            |||||||||||||||||||||||||||
Db  968  AAGAAGCTTGAGGTGAACAAGAGCAGTATCAAGTACGAGAATCGATTCTTG  1027

Qy  321  PheArgLysGlyGluValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArg  340
            |||||||||||||||||||||||||||||||
Db 1028  TTTCGGAAAGGAGAAGTGAGTGATTGGGTTAACTATTTGTCACCTTCACAAGTGGAAAGA  1087

Qy  341  LeuSerAlaLeuValAspAspLysLeuAlaGlyGlySerGlyLeuThrPheArgLeuSer  359
            |||||||||||||||||||||||||||||||
Db 1088  TTGTCAGCCCTTAGTGGATGACAAGTTAGGTGGTGGAAGTGGGCTCACTTTCAGGTTGAGC  1144
```

RESULT 10
LOCUS   AX080460   1041 bp   DNA   linear   PAT 22-FEB-2001
DEFINITION Sequence 3 from Patent WO0102589.
ACCESSION   AX080460
VERSION   AX080460.1 GI:13159886
KEYWORDS
SOURCE   Arabidopsis thaliana (thale cress)
  ORGANISM Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Varin,L. and Gidda,S.
  TITLE   Methods, compositions and genetic sequences for modulating
            flowering in plants, and plants genetically modified to flower
            early and tardily
  JOURNAL   Patent: WO 0102589-A 3 11-JAN-2001;
            Varin, Luc (CA) ; Gidda, Satinder (CA)
FEATURES   Location/Qualifiers
     source   1..1041
                     /organism="Arabidopsis thaliana"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3702"
ORIGIN

US-10-019-931-3 (1-359) x AX080460 (1-1041)

Alignment Scores:
Pred. No.:   11
Score: 2.08e-126
Score: 1553.00   Length: 1041
Percent Identity: 92.31%   Matches: 299
Percent Similarity: 85.19%   Conservative: 25
Best Local Similarity: 81.82%   Mismatches: 21
Query Match: 6   Indels: 6
DB: Gaps: 3

Qy   11 MetAlaIleProSerPheSerMetCysHisGlyLeuLeuLysGlyLeuGlyLysThr 30
         ||||:::|||| |||| | |||||
Db   1 ATGGCGATCCCAAGTTTCTCCATGTGTCACAAGCCCGAGCTCCTTAAGGGAAGGCAAAGC 60

Qy   31 ArgAspValProLysGluGluAspGlyLeuSerCysGluPheGlnGluMetLeu 50
         |||:||:|||
Db   61 --------GAAGGCCAAGAAGAAGAAGGGCTTAAGCTACGAGTTCCAAGAAGAGATGTTG 108

Qy   51 AspSerLeuProLysGlyLeuArgGlyTrpArgThrArgTyrLeuTyrLeuPheGlnGlyPhe 70
         |||||||||||||||||||||||||||||||||
Db   109 GACTTCTTCCTAAGGGAGAGGAGAGGACCGGAGAAATCGTTACTTAATTCCAAGGGTTT 168

Qy   223 TyrCysArgGly---ValIleGlyPheGlyProTrpGluHisMetLeuGlyTyrTrp 241
Db   693 TATTGCCGCGGGATTACTGATCGGATTTGGCCCGTTTGGGAACACATGTTGGGATACTGG 752

Qy   242 ArgGluSerLeuLysArgProGluLysValPheProPheLeuArgTyrGluAspPheAsp 261
Db   753 AGAGAGAGCTTGAAGAGGCCAGAGAAAGTCTTATTTTAAAGTACGAGGATCTCAAAGAA 812

Qy   262 AspIleGluThrAsnLeuLysArgLeuAlaThrArgLeuProPheThrGluGlu 281
Db   813 GACATCGAGACCAACTTGAAGAGGCTAGCAACTTCCTTTCACCGAAGAA 872

Qy   282 GluGluArgLysGlyValValLysAlaLeuAlaLeuCysSerPheGluAsnLeuLys 301
Db   873 GAGGAACAAAGGGGAGTTGTGAAAGCTATCGCTCTGTGTAGCTTTGAGAATCTGAAG 932

Qy   302 LysLeuGluValAsnLysSerAsnLysSerIleLeuGluCysValPheAsnPheLeuPhe 321
Db   933 AAGTTGGAGGTGAACAAGTCAAACAAATTGAATTGTGTGTTTGTGAACTTTCTTGTTT 992

Qy   322 ArgLeuGlyGluValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArgLeu 341
Db   993 AGGGAAAGGGAGGAAGTGAGTGATTGGGTTAACTATTTGTCGCCATCACAAGTGGAAAGAT 1052

Qy   342 SerAlaLeuValAspAspLysLeuGlyGlySerGlyLeuLeuThrPheArgLeuSer 359
Db 1053 TCACCCTTAGTCGGATGACAAGTTACTGGTCTCACTTTCAGATTGAGT 1106

RESULT 11
LOCUS   BT006241   1044 bp   mRNA   linear   PLN 17-APR-2003
DEFINITION Arabidopsis thaliana At5g07000 mRNA, complete cds.
ACCESSION   BT006241
VERSION   BT006241.1 GI:30017312
KEYWORDS   FLI_cDNA.
SOURCE   Arabidopsis thaliana (thale cress)
  ORGANISM Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Qy   349 LeuGlyLeuGlySerGlyLeuLeuThrPheArgLeuSer 359
Db 1009 TTAGGCTGGATCTGGGTCTCACTTTCAGATTGAGT 1041

```
REFERENCE   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
  AUTHORS   1 (bases 1 to 1044)
            Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P.,
            Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Heuan,V.W.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Tang,C.C., Wu,H.C., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K.,
            Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
            Ecker,J.R.
  TITLE     Arabidopsis ORF clones
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1044)
  AUTHORS   Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Bowser,L., Carninci,P.,
            Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Heuan,V.W.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K.,
            Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
            Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-APR-2003) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA ; 'RIKEN
            Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.
            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim,C.J.,
            Chen,H., Cheuk,R., Shinn,P., Bowser,L., Chan,M.M., Chang,C.H.,
            Dale,J.M., Heuan,V.W., Jones,T., Karlin-Neumann,G., Lam,B.,
            Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
            Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
            Yamada,K., Yu,G., Yuan,S., Davis,R.W., Theologis,A. and
            Ecker,J.R.
            Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES            Location/Qualifiers
     source          1..1044
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="U60156"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51"
     CDS             1..1044
                     /codon_start=1
                     /product="At5g07000"
                     /note="putative steroid sulfotransferase"
                     /protein_id="AAP12890.1"
                     /db_xref="GI:30017313"
                     /translation="MAIFSFSMCHKPELLKEGKSEGQEEEGLSYEFQEMLDSLPKERG
                     RRNRYLYLFQGFRCQAKEIQAITSFQKHFQSLPDDVVLATIPKSGTTWLKALFTLILT
                     KDSVENPSVKVVVLCRRPFDTFISMVHINNTISESVSAVLLDEAFLYCRGLLIGF
                     RHRFDPVSSSSSSDHPLLISNPHDLVPFFEYKLYANGNVPDLSGLLASPRTPATHVPFGA
                     GPWEHMLGYWRSLKRPEKVLFLKYEDLKEDIETNLKLASFLGLPFTEEEOKGVV
                     KAIADLCSFENLKKLEVNKSSKLIQNYENRFLFRKGEVSDLVNYLSPSQVERLSALVD
                     DKLAGSGITFRLS"
ORIGIN
```

Alignment Scores:
Pred. No.:   f
Score: 2.09e-126
Percent Similarity: 92.31%

```
                                          Length:        1044
                                          Matches:       299
                                          Conservative:  25
```

US-10-019-931-3 (1-359) x BT006241 (1-1044)

```
Best Local Similarity:   85.19%          Mismatches:  21
Query Match:             81.82%          Indels:       6
DB:                      15              Gaps:         3
```

```
Qy    1  MetAlaIleProSerPheSerMetCysHisLysLeuLeuLeuLysGluGlyLysThr  30
         ||||||||||||||||||||||||||||||||| ||: ||||||||||||||| |||
Db    1  ATGGCCGGATCCCCAAGTTTCTCCATGTGTCACAAGCCCCAGCTCCTTAAGGAAGGCAAAGC  60

Qy   31  ArgArgAspValProLysAlaGluGlyLeuAspGluGlyLeuSerCysGluPheGlnGlyMetLeu  50
         :||   |||||||||||||||||  ||  :||
Db   61  GAAGGCCGAAGAAGAAGAAGAGGGCTAAGCTACGAGTTCCAAGAGATGTTG  108

Qy   51  AspSerLeuProLysGlyLeuArgGlyTrpArgThrArgTyrLeuPheGlnGlyPhe  70
         ||||||||||||||:||||::||||||||||:::|:||||||||||
Db  109  GACTCTTTGCCTAAGGAAGAAGAGGACGGAGAAATCGTTACCTTAATTCCAAGGGTTT  168

Qy   71  TrpCysGlnAlaLysGluIleGlnAlaIleMetSerPheGlnLysHisPheGlnSerLeu  90
         |||||||||||||||||||| ||||||||||||||||||||||||||
Db  169  CGGTGCCAAGCTAAGGAAGATTCAAGCTATCACGTCTTTCCAAAAACATTTCAGTCCCTT  228

Qy   91  GluAsnAspValValLeuAlaThrIleProLysSerGlyThrThrTrpLeuLysAlaLeu  110
         ::||||||||||| ||||||||||| ||||||||||||:|::
Db  229  CCAGACGACGTTGTTCTCGCCACCATACCTAAAATCTGGCACAACTTCGGTTAAAAGCTTTA  288

Qy  111  ThrPheThrIleLeuAsnArgHisArgPheAspProVal---AlaSerSerThrAsnHis  129
         |||||||||||::: ||||:::  :|||| |:||:|  ::||||||::::|||
Db  289  ACTTTCACCATCTTAACCCGTCACATCCGGTTTCCTCATCAAGTTCCGACCAC  348

Qy  130  ProLeuPheThrSerAsnProHisAspLeuValProPhePheGluTyrLysLeuTyrAla  149
         ||||:|||||||||||||||||:::||||||||||||||||||||||:|||
Db  349  CCTCTTCACATCCAACCCCACGACCTCGTACCTTTCCTCGAGTACAAGCTTTACGGC  408

Qy  150  AsnGlyAsnValProAspLeuSerGlyLeuLeuAlaSerProArgThrProAlaThrHisLeu  169
         ||||||:||||||||||||:::|:|||| ||||||||||||||||||:::
Db  409  AACGGAAATGTTCCCGATCTCTCCGGGTCTACTAGGACCACCACGTA  468

Qy  170  ProPheGlySerLeuLysGluThrIleGluLysProGlyValLysValValValLeuCys  189
         ||||||||||:::||||||||||||:::||||||||||||||||||||:||:
Db  469  CCGTTCGGTGCCCTTAAGGAATTCGGTCGAGAATCCAGTGGTGAAGGTTGTGTTACCCTGTGC  528

Qy  190  ArgArgProPheAspThrPheIleSerMetValHisIleTyrThrAsnAsnIleLysSerGlu  209
         ||||||||||:||||||||||||||||||||||||:::|||||||||||
Db  529  CGGAACCCGTTTGACACATTCATCTCCATGGCATTACTCCATGGCAATCAACAACATCACTTCCGAG  588

Qy  210  SerValSerProValLeuLeuLeuAspGluAlaPheAspLeuTyrCysArgGlyLeu---Valile  228
         ||||||||| ||||||||||||||||:::   |:||||:|| ||:::|::||
Db  589  TCAGTGAGCGCGCAGTCTTGCTAGACGAAGCTTTGACCTATAATTGCCGGGAATTACTGATC  648

Qy  229  GlyPheArgProPheTrpGluHisMetLeuGlyTyrTrpArgSerLeuLysArgPro  248
         |||||||||||||||||||:::|||||||||||::||||||
Db  649  GGATTCGGCCCGTTTTGGGAGCACATGTTGGGATACTGGAGAAGCCTGAAGAGGCCA  708

Qy  249  GluLysValPheLeuLeuLysTyrGluAspLeuLysGluAspIleGluThrAsnLeuLys  268
         ||||||| ||||||||||||:::|||||||||:|||||||||||||||||||
Db  709  GAGAAAGTCTTTATTTTAAAGTACGAGGATCTCAAAGAAGACATCGAGACCAACTTGAAG  768

Qy  269  ArgLeuAlaThrPheLeuProPheThrGluGluGluGlnLysGlyValVal  288
         ||||||||||| ||::: ||||||||||| |||:::||||
Db  769  AAGCTAGCAACGTTTCTTAGGACTTCCTTTACCGAAGAAGAGGAACAAAAAGGAGGTGTG  828

Qy  289  LysAlaIleAlaAspLeuCysSerPheGluAsnLeuLysLysLeuGluValAsnLysSer  308
         ||||||||||||:::|||||||||||||||||||||||||||||||||||||||
Db  829  AAAGCTATCGCTGATCTGTGTTCGTTTGAGAATCTGAAGAAGTTGGAGGTGAACAAGTCA  888

Qy  309  AsnLysSerLysLeuIleGlnAsnAsnArgPheLeuPheArgLysGlyGluValSerAsp  328
         :::||| |||||||||||| ||||||||||||:::|||||:|| ||||||
Db  889  AGCAAATTGATCCAGAACTATGAGAACCGGTTCTTGTTTAGGAAAGGAGAAGTGAGTGAT  948

Qy  329  TrpValAsnTyrLeuSerProSerGlnValGluArgLeuSerAlaLeuValAspAspLys  348
         :::|||| ||||||||||||||||||::|||||||||||:::|||:|||||||
Db  949  TTGGTTAACTATTTGTCGCCATCACAAGTGGAAAGATTGTCAGCCCTTAGTGGATGACAAG  1008
```

Qy   349 LeuGlyGlySerGlyLeuThrPheArgLeuSer 359
            |||||||||||||||||||||||||||||||||
Db   1009 TTAGCTGGATCTGGTCTCACTTCAGATTGAGT 1041

RESULT 12
Db   CQ832235
LOCUS   CQ832235   1023 bp   DNA   linear   PAT 29-JUL-2004
DEFINITION Sequence 2 from Patent WO2004057000.
ACCESSION   CQ832235
VERSION   CQ832235.1 GI:50831891
KEYWORDS   .
SOURCE   Arabidopsis thaliana (thale cress)
  ORGANISM Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Varin,L.
  TITLE   Methods and compositions for modulating tuber formation in
            tuber-producing plants
  JOURNAL   Patent: WO 2004057000-A 2 08-JUL-2004;
            Florigys Inc. (CA)
FEATURES   Location/Qualifiers
     source   1..1023
                     /organism="Arabidopsis thaliana"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3702"

ORIGIN

Alignment Scores:
Pred. No.:
Score: 1.92e-123   Length: 1023
                         1519.00   Matches: 292
Percent Similarity: 92.15%   Conservative: 25
Best Local Similarity: 84.88%   Mismatches: 21
Query Match: 80.03%   Indels: 6
DB: 6   Gaps: 3

US-10-019-931-3 (1-359) x CQ832235 (1-1023)

Qy   18 MetCysHisHisLysLeuGluLeuLeuLysGluGlyLysThrArgAspValProLysAlaGlu 37
            |||||||||||||||||||||||||||||||||||||:||
Db   1 ATGTGTCACAAGCCCGAGCTCCTTAAGGAAGGCAAAAGC----------GAAGGCCAA 48

Qy   38 GluAspGlyGlyLeuSerCysGluPheGlnGlnMetLeuAspSerLeuProLysGluArg 57
            |||||||||:||||||||||||||||||:::||||||||
Db   49 GAAGAAGGAGGGCTAAGCTACACGAGTTCCAAGAGAGTTGGACACTCTTCCTAAGGAGAGA 108

Qy   58 GlyTrpArgThrArgTyrLeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGluIle 77
            |||||||||||||||||||||||||||||||||||||
Db   109 GGACGGAGGACAAATCGTTACTTTACTTTACTTCAAGGGTTTCGGTCGCCAAGCTAAGGAGATT 168

Qy   78 GlnAlaIleMetSerPheGlnLysHisPheGlnSerLeuGluAsnAspValValLeuAla 97
            |||||||||||||:::|||||||||||||||||||||||||
Db   169 CAAGCTATCACGTCTTTCCAAAAACATTTTCAGTCCCTTCCAGACAGACGTTGTCCTCGCC 228

Qy   98 ThrArgProLysSerGlyThrThrTrpLeuLysAlaLeuThrPheThrThrIleLeuAsnArg 117
            ||||:|||||||||||||||||:||||||||||||||:|
Db   229 ACCATACCTAAAATCTGGCACAACCTGGTTAAAAGCTTTAACTTTCACCATCCTTACCCGT 288

Qy   118 HisArgPheAspProVal---AlaSerSerThrAsnHisProLeuPheThrSerAsnPro 136
            ::|||||||||||||| ::|||:|||||:||||||||||:||:||||||
Db   289 CATCGGTTTGATCCGGTTTCCTCATCCAAGTTCCACCACCCTCTTCTCACATCCAACCCT 348

Qy   137 HisAspLeuValProPhePheGluTyrArgLysLeuTyrAlaAsnGlyAspValProAspLeu 156
            ||||||||||||||||||||||||:||||||||||||||||:||||||
Db   349 CACGACCTCGTACCTTTCTTCGAGTACAAGAGCTTTACGCAAACGGAAAGTTCCCGGATCTC 408

Qy   157 SerGlyLeuAlaSerProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGlu 176
            ||||||||||||:||||||||||||||||||||:::|||||||:::
Db   409 TCCGGTCTAGCCAGTCCAAGAACAATTCGGAATGTTCGGTCGCCCCTTAAGGAT 468

Qy   177 ThrIleGluLysProGlyValLysValValTyrLeuCysArgAsnProPheAspThrPhe 196

RESULT 14

```
Qy     18 MetCysHisLysLeuGluLeuLeuGluGluGlyLysThrArgArgAspValProLysAlaGlu 37
          |||||||||||||||||||||||||||||||||||||||||||||||::::
Db      1 AtGTGTCACAAGCCCGAGCTCCTTAAGGAAGGCAAAAGC--------------GAAGGCCAA 48

Qy     38 GluAspGluGlyGlyLeuSerCysGluPheGlnGlnMetLeuAspSerLeuProLysGluArg 57
          ||||::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 GAAGAAGAAGGGCTAAGCTACGAGTTCCAAGAGATGTTGGACTCTCTTCCTTAAGGAGAGA 108

Qy     58 GlyTrpArgThrArgThrTyrLeuPheTrpCysGlnAlaLysGlyIleIle 77
          ||||||||||||||||||||||||||||||||||||||||||||::|
Db    109 GGACCGGAGAAATCGTTACCTTACTTAATTCCAAGAGGTTTCCGGTGCCAAGCTAAGGATT 168

Qy     78 GlnAlaIleMetSerPheGlnLysHisSerPheGlnLeuAsnAspValValLeuAla 97
          |||||||||||||||||||::||||||||||||||||||||||||||||
Db    169 CAAGCTATCACGTCTTTCAAAAACATTTCAGTCCCTTCCAGACCGACGTTGTCCTCCGCC 228

Qy     98 ThrIleProLysSerGlyThrThrTrpLeuLysAlaLeuThrPheThrThrIleLeuAsnArg 117
          |||:::||||||||||||||||||||||||:::||||||||||||||||||||||||
Db    229 ACCATACCTAAAATCTGGCACAACCTGGTTAAAAGCTTTAACTTTCACCATCCTTACCCGT 288

Qy    118 HisArgPheAspProVal---AlaSerSerThrAsnHisProLeuPheThrSerAsnPro 136
          ||||||||||||||||||   :::||||||||||||||||||||::||||
Db    289 CATCCGTTTGATCCGGTTTCCTCATCCAAGTTCCGACCACCCTTCTCACATCCCAACCCT 348

Qy    137 HisAlaAspLeuValProPhePheGluTyrLysLeuSerLeuTyrAlaAsnGlyAspValProLeu 156
          |||||||||||||||||||||||||||||||::||||::|||
Db    349 CACGACCTCGTACCTTTCTTCGAGTACAAGCTTTACGCCAACGGAAATGTTCCCGGATCTC 408

Qy    157 SerGlyTyrLeuAlaSerProArgThrThrPhePheAlaThrHisLeuProPheGlySerLeuLysGlu 176
          |||||||||||||||||::||||||||||||||||||::||||
Db    409 TCGGGCTACCTAGCCAGTCCAAGAACAATTCGCAACCCACGTACCGTTCGGTCGCCTTAGGAT 468

Qy    177 ThrIleGluLysProGlyValLysValValTyrLeuCysArgArgAsnProPheAspThrPhe 196
          :::::||||||||||||||||||||||||||||||||::||
Db    469 TCGGTCGAGAATCCCAGTGTGAAGGTTGTGTACCGTGCCGGACCCGTTTGACACATTC 528

Qy    197 IleSerSerTrpHisTyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeu 216
          |||::::|||||||||||||||||||||||||||||||||||||||||||||||
Db    529 ATCTCCAGTTGGCATTACATCACAACAATCACTCCGAGTCAGTGAGGCCAGTCTTGCCTA 588

Qy    217 AspGlnAlaAlaPheAspLeuTyrCysArgGly---ValIleGlyPheGlyTyrTrpGlu 235
          |||:::|||||||||||||||||||||   :::|||||||||||||||||||
Db    589 GACCAGGAAGCTTTTGATCTATATTGCCGGGGGATTACTGATCGGATTTGGCCCGTTTGGGAA 648

Qy    236 HisAlaMetLeuGlyTyrThrTyrArgGluSerArgProGlyLysValPhePheLeuArg 255
          ||||||||||||||||||::||||||||||||||||||||||||||::
Db    649 CACATGTTGGGATACTGGAGAGAGGCGCAGAGAAGTCTTTATTTTTAAAG 708

Qy    256 TyrGluAspLeuLysLysAspAsnProIleGluGluThrArgAsnLeuTyrArgLeuAlaThrPheLeuGlu 275
          ||||||||||||||||:::|||||||||||||||||||||::||||
Db    709 TACGAGGATCTCAAAAGAACAATCGAGAGACCACTGAACAAGTTCTAGGAGTTTCTTAGGA 768

Qy    276 LeuProPheThrGlyGluGlyLysArgGlyLysValValAlaIleAlaGluLeuLeuCys 295
          ||||||||||||||||||||||||||||||||||||||||:::||::||:::
Db    769 CTTCCTTTCACCGGAGAGAAGGAGGCCCAGAGAAAGTCTATCGGTCGATCTGTGT 828

Qy    296 SerPheGluAsnLeuLysLeuGluValAlaSerSerIleLysSerAsnPhe 315
          ||||||||||||||||||||||||||||||:::||::
Db    829 AGCTTTGAGAATCTGAAGAAGTTGAGGGTGAACAAATTGACCCAGAACTAT 888

Qy    316 GluAsnArgPheLeuProPheArgLysGlyValValSerAspTrpValAlaSerTyrLeuSerPro 335
          |||||||||||||||||||||||||||||||||||::||||||
Db    889 GAGAACCCGGTTCTTCGTTTAGGAAAGGAGTGAGTGTATCGGATTGGTTAACTATTGTCGGCCA 948

Qy    336 SerGlnValGlyLeuSerAlaLeuValAspAlaProLysGlyLeuGlySerGlyLeuThr 355
          ||||||||||||||||||||||||||||||||||::||||||||||||||||
Db    949 TCACAAGTGGAAAGATTGTCAGCCTTAGTGGATGACAAGTTAGTGGATCTGGTCTCACT 1008

Qy    356 PheArgLeuSer 359
          ||||||||||||
Db   1009 TTCAGATTGAGT 1020
```

```
Alignment Scores:
Pred. No.:

Score:                  1076.50      Matches:        196
Percent Similarity:     74.43%       Conservative:    66
Best Local Similarity:  55.68%       Mismatches:      77
                        2.2e-82      Length:      115071
```

```
Query Match:          56.72%    Indels:    13
DB:                   14         Gaps:      4

US-10-019-931-3 (1-359) x AC135566 (1-115071)
```

```
Qy          9  IleProMetAlaIleProSerPheSerMetCysHisLysLeuGluLeuLysGluGlyGly  28
               |||||       ||:||
Db      31498  ATTTCAATGTCTTCAACAAAAATTGCCAAAATCCAC------------------------  31533

Qy         29  LysThrArgAspValProLysAlaGluGluGluAspGluGly---LeuSerCysGluPheGln  47
                         |:::                                 |||
Db      31534  ---ACAATGGATATAGATAGAATCCATTGAGAACAAGGTCAACTTAGCCAAGAAAACAAG  31590

Qy         48  GluMetLeuAspSerLeuProLysGluArgGlyTyrThrArgTyrLeuTyrLeuPhe  67
               ::::::::::     |||:||:::                     ||:||:|||
Db      31591  AAACTAATTCTTCTTTCCAAGAGAGAAAGGTTGGAGAACACCTTATATTATCTTATTC  31650

Qy         68  GlnGlyPheTrpLysGlnAlaLysGluIleGlnAlaIleMetSerPheGlyLysHisPhe  87
               |||:|||||||                                    ||||:|||
Db      31651  CAAGGATTTTGGTGCCAACCAGCTGAAATCCAGCCATAACCACTTTCCAAAAGGCATTTC  31710

Qy         88  GlnSerLeuGluAsnAspValValLeuValAlaThrIleProLysSerGlyThrThrTrpLeu  107
               |||:||                    |||                    |||
Db      31711  CAAGCTAAAGAAAAGTGATGTTTTTGTTGGTTGCAACCGTGACCAAAATCAGGCACAACTTGGTTA  31770

Qy        108  LysAlaLeuThrPheThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThr  127
               |||:|||||:||||                    |::
Db      31771  AAAGCTCTTACCTTTACCTATCGCCATTATGAAATCGCCAA------AACCATTCATTTCATCCAAA  31824

Qy        128  AsnHisProLeuPheThrSerAsnProHisAspLeuValProPhePheGluTyrLysLeu  147
               |||||||||:|||                    :::                  ||
Db      31825  AACCATCCTTTGCTTAGTTTCAATCCACATGATCTTGTTCCTTTCATTGAGTATACAGTT  31884

Qy        148  TyrAlaAsnGlyAsp---ValProAspLeuSerGlyLeuValAlaSerProArgThrPheAla  166
               ::                        |||                  ||||
Db      31885  TATGGTAAGCATGACACAATTCCTGATTTCTCTAAAATTCATGAGCCTAGACTTTTTGGT  31944

Qy        167  ThrHisLeuProPheGlySerLeuLysGluThrIleGlyLysProGlyLysLeuValVal  186
               :||||     ||||
Db      31945  ACACATTAATTCCTTTGGACTCAGTCCAATTGCCAAATAGTTA  32004

Qy        187  TyrLeuCysArgAsnProPheAspThrIleLeuSerSerTrpHisTyrThrArgAsnIle  206
               ||:||                             :::
Db      32005  TATATTTGTAGGAACCCTTTTGATACTTTCATCTCTTCCTGGCATTATACAAGAAATT  32064

Qy        207  LysSerGluSerValSerProValLeuLeuAspGlnAlaPheAspLeuTyrCysSerArgGly  226
               |||||||                                   ||||
Db      32065  AAGCCACCTTCTTTGCCTACATTGAATTTAGAGGAGGCTTTTGAAATGTTATTGTAATGGT  32124

Qy        227  ValIleGlyPheProThrPheTrpGluHisMetLeuGlyTyrTrpTyrThrArgGluSerLeuLys  246
               :::||||||                           |||
Db      32125  TTAGTTGGGTTTGGGCCCATTTGGAAATCACAGTTAGGTTTATTGGAATGAGAGCAAAGAG  32184

Qy        247  ArgProGluLysValProPhePheLeuArgTyrGluAspIleGluLysLysSerAspIleGluThrArgAsn  266
                                                  ||||
Db      32185  AGACCAAAAAAAATGTTCTTTTTTTGAAGTATGAGGAGGATTTAAAATTTCAC  32244

Qy        267  LeuLeuSerArgLeuAlaThrPheLeuGluLeuProPheThrGluGluGluValLysGly  286
                                                  ||||
Db      32245  TTGAAAAAAATGGCTGAGTTTTTTGATCGTCCTTTTACAATTGGAGGAGAAAGTGAGGGT  32304

Qy        287  ValValLysAlaIleAlaGluLeuCysSerPheGluAsnLeuLysLeuGluValAsn  306
               |||:||                             ||:
Db      32305  GTGATTGAAAATATAATAAAAGTGTCGTAGTTTGAGAAAATGAAGAGGGTTTGGAGATAAAAT  32364

Qy        307  LysSerAsnLysSerIleLeuLysAsnPheLeuAsnArgPheLeuArgLysGlyGlyLysLysVal  326
               :::                                ||||
Db      32365  AATACAGGAACATTTGGAAGGAACATTTGAGAAATAAGTTCTTTGTTTTGAGAAAAGGTGAAGTT  32424

Qy        327  SerAspTrpValAsnTyrLeuSerProSerGlnValGlyLeuArgLeuSerAlaLeuValAsp  346
               ||||                               ||:|||
Db      32425  GGTGATTGGAGTAATATCTTTCAATGGTTGAAAAAAAATTGTCCAAGGTCATTGAA  32484
```

```
Qy        347  AspLysGluGlyGlySerGlyLeuThrPheArgGlu  358
               |||||||||||||||||||||||||||||
Db      32485  GAGAGTTAGGTGGATCTGATCTCAATTTTAGAGTG  32520
```

```
RESULT 15
Db    32485. GAGAGTTAGGTGGATCTGATCTCAATTTTAGAGTG

AP008111
LOCUS       AP008111      100108 bp    DNA    linear    HTG 28-DEC-2004
DEFINITION  Lotus corniculatus var. japonicus clone LjT05D15, *** SEQUENCING IN
            PROGRESS ***, 20 unordered pieces.
ACCESSION   AP008111
VERSION     AP008111.1  GI:56806418
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Lotus corniculatus var. japonicus (Lotus japonicus)
  ORGANISM  Lotus corniculatus var. japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1
  AUTHORS   Kaneko,T., Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     Structural Analysis of a Lotus japonicus Genome. XI. Sequence
            Features and Mapping of Nine hundred twenty-one TAC Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 100108)
  AUTHORS   Sato,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-OCT-2004) Shusei Sato, Kazusa DNA Research Institute,
            Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
            Chiba, 292-0818, Japan (E-mail:ssato@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337),
            Fax:81-438-52-3934)
COMMENT     NOTE: This is a 'working draft' sequence. It currently
            consists of 20 contigs. The true order of the pieces
            is not known and their order in this sequence record is
            arbitrary. Gaps between the contigs are represented as
            runs of N, but the exact sizes of the gaps are unknown.
            This record will be updated with the finished sequence
            as soon as it is available and the accession number will
            be preserved.

            1            5202:  contig of 5202 bp in length
            5203         5302:  gap of unknown length
            5303         6566:  contig of 1264 bp in length
            6567         6666:  gap of unknown length
            6667         7760:  contig of 1094 bp in length
            7761         7860:  gap of unknown length
            7861         8624:  contig of 764 bp in length
            8625         8724:  gap of unknown length
            8725         9607:  contig of 883 bp in length
            9608         9707:  gap of unknown length
            9708        11541:  contig of 1834 bp in length
           11542        11641:  gap of unknown length
           11642        13795:  contig of 2154 bp in length
           13796        13895:  gap of unknown length
           13896        15742:  contig of 1847 bp in length
           15743        15842:  gap of unknown length
           15843        19031:  contig of 3189 bp in length
           19032        19131:  gap of unknown length
           19132        21810:  contig of 2679 bp in length
           21811        21910:  gap of unknown length
           21911        24667:  contig of 2757 bp in length
           24668        24767:  gap of unknown length
           24768        30306:  contig of 5539 bp in length
           30307        30406:  gap of unknown length
           30407        36585:  contig of 6179 bp in length
           36586        36685:  gap of unknown length
           36686        42389:  contig of 5704 bp in length
           42390        42489:  gap of unknown length
           42490        50891:  contig of 8402 bp in length
           50892        50991:  gap of unknown length
           50992        60141:  contig of 9150 bp in length
           60142        60241:  gap of unknown length
           60242        69078:  contig of 8837 bp in length
           69079        69178:  gap of unknown length
```

```
       * 69179  79488: contig of 10310 bp in length
       * 79489  79588: gap of unknown length
       *         98477: contig of 18889 bp in length
       * 98478  98577: gap of unknown length
       * 98578 100108: contig of 1531 bp in length.

FEATURES             Location/Qualifiers
     source          1..100108
                     /organism="Lotus corniculatus var. japonicus"
                     /mol_type="genomic DNA"
                     /variety="japonicus"
                     /db_xref="taxon:34305"
                     /clone="LjT05D15"
                     /clone_lib="LjT library"
                     /note="TAC clone:TM1490, synonym:Lotus japonicus"
     gap             5203..5302
                     /estimated_length=unknown
     gap             6567..6666
                     /estimated_length=unknown
     gap             7761..7860
                     /estimated_length=unknown
     gap             8625..8724
                     /estimated_length=unknown
     gap             9608..9707
                     /estimated_length=unknown
     gap             11542..11641
                     /estimated_length=unknown
     gap             13796..13895
                     /estimated_length=unknown
     gap             15743..15842
                     /estimated_length=unknown
     gap             19032..19131
                     /estimated_length=unknown
     gap             21811..21910
                     /estimated_length=unknown
     gap             24668..24767
                     /estimated_length=unknown
     gap             30307..30406
                     /estimated_length=unknown
     gap             35586..35685
                     /estimated_length=unknown
     gap             42390..42489
                     /estimated_length=unknown
     gap             50892..50991
                     /estimated_length=unknown
     gap             60142..60241
                     /estimated_length=unknown
     gap             69079..69178
                     /estimated_length=unknown
     gap             79489..79588
                     /estimated_length=unknown
     gap             98478..98577
                     /estimated_length=unknown
ORIGIN
```

Alignment Scores:

Pred. No.:   9
Score: 7.34e-78
Score: 1024.00
Percent Similarity: 73.09%
Best Local Similarity: 55.81%
Query Match: 53.95%
DB: 14

Length: 100108
Matches: 197
Conservative: 61
Mismatches: 85
Indels: 11
Gaps: 4

US-10-019-931-3 (1-359) x AP008111 (1-100108)

```
Qy     9 IleProMetAlaIleProSerPheSerMetCysHisLysLeuLeuLysGluGluGly 28
           ||||||  |  ||| |||  ||| | |||||   |     |
Db 60863 ATTCCAATGGACACCAACAGCTCCAACCATGTTTGCAAAATTCCAT---     60907

Qy    29 LysThrArgAspValProLysGluAlaGlyLeuSerCysGluPheGlnGlu 48
           :::                  |||||  ||||
Db 60908 ---TCAAATGGTGATGAGGGAGCAATTGATCAAGAGAACTTAGTGAAGAAATGAAGGAT 60964

Qy    49 MetLeuAspSerLeuProLysGlyLeuArgGlyTrpArgThrArgTyrLeuTyrLeuPheGln 68

Db 60965 CTCATTCTCTCTCCAGA-GAGAGAGGTTGGAGAACCAAATATCTATATCTCCAA     61023

Qy    69 GlyPheTrpCysGlnAlaLysGlyIleGlnAlaIleMetSerPheGlnLysHisGln 88
Db 61024 GGCTTTTGGGTGTGCCAACCAGCTGAAAATCCAAGCCATTAAGCACTTTCAGAAGCATCAA 61083

Qy    89 SerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThrTrpLeuLys 108
Db 61084 GCCAAGACAGTGATGTTGTCTTTGGTTGCCACAGTTCCAAAATCAGGGTGGTTAAAG     61143

Db 61144 GCACTAACCTTTGCCATTGTAAACCGCCAGAAACAT----TTAACCTTCCAGAGC     61197
Qy   109 AlaLeuThrPheAlaThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThrAsn 128

Qy   129 HisProLeuPheThrSerAsnProHisAspLeuValProPhePheGluTyrLysLeuTyr 148
Db 61198 CACCCTTTATTAAATTCCAATCCCATGACCTTGTCCTTTCAAATACTACACTGTTTAT     61257

Qy   149 AlaAsnGlyAsp---ValProAspLeuSerGlyLeuAlaSerProArgThrPheAlaThr 167
Db 61258 GGGAAACAATGACAAGGTTCCTGATTTATCCAAAATCCCTGACCCAAGGGCTTTTGGAACC 61317

Qy   168 HisLeuProPheGlySerLeuLysGluThrIleGluGluLys---ProGlyValLysVal 186
Db 61318 CACATTCCATTTTCTTCTCTGCCCAAATTCAATCACCAAAAGGTCCAAATTGCAAGAATTGTT 61377

Qy   187 TyrLeuCysArgAsnProPheAspThrPheIleSerSerSerTrpHisTyrThrAsnAsnIle 206
Db 61378 TATATCTGCAGGAACCCTTTTGACACTTTCATCTCTTCTTCGTGGCATTACACTAACAACAATC 61437

Qy   207 LysSerGluSerValSerProValLeuLeuAspHisGlnIleAlaPheAspHisLeuTyrCysArgGly 226
Db 61438 AAACCACCTTCAATTACCCACATTGATGCTTGAGGAGGCTTTGACAGTTTTGTAATGGG 61497

Qy   227 ValIleGlyPheGlyProPheTrpGluHisMetLeuGlyTyrTrpArgGluSerLeuLys 246
Db 61498 GTGGTTGGGTTTGGCCCATTTTGGGAATCACAATGTTGGGTTAATGGAGAGCATTGAG     61557

Qy   247 ArgProGlyLysValPhePheLeuArgTyrGluAspLeuGluThrThr 266
Db 61558 AGACCAAACAAGGTTCTTTTTCTTAGAGTGGCAATGAAAGAAGATCCTGTTTTTCAC     61617

Qy   267 LeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGluGluGluLysGlyGly 286
Db 61618 TTGAAAAGAGATGCTAGTTGTTGCTTCTTCCTTGGAGGAGAAGGGGGGT     61677

Qy   287 ValValLysAlaIleAlaGluLeuCysSerPheGluAsnLeuLysLysLeuGluValAsn 306
Db 61678 GTAAATTGAGAACATAATTGATTTGGTGTTAGCTTTGAGAAGATGAAGAAATTGAGGTAAAT 61737

Qy   307 LysSerAsnLysSerIleLysAsnPheGluAsnArgPheLeuPheArgLysGlyGlyVal 326
Db 61738 AAGTCTTGGAACATTGCCAGGAAATTCAAGAACAAGTTCTTGTTCAGGAGGCCATGGGGT     61797

Qy   327 SerAspTrpValAsnThrTyrLeuSerProSerGlnValGlyArgLeuSerAlaLeuValAsp 346
Db 61798 GGGGATTTTGTGAATATCACTTTCATGGTGGAGAAATTATCCAAAAGTCATGGAA     61857

Qy   347 AspLysLeuGlyLysSerGlyLeuThrPheArgLeuSer 359
Db 61858 GAGAAGCTAGGTGGGGTGGGTTGTTGTCATTTAGAGTGAAC     61896
```

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: December 23, 2005, 16:54:45 ; Search time 604 Seconds
                                    (without alignments)
                                    3961.299 Million cell updates/sec

Title: US-10-019-931-3
Perfect score: 1898
Sequence: 1 MATSSMKSIPMAIPSFSMCH.........RLSALVDDKLGGSGLTFRLS 359

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop   6.0 , Fgapext 7.0
Delop   6.0 , Delext   7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10019931/runat_16122005_165218_18429/app_query.fasta_1.519
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10019931 @CGN 1_1 1096 @runat 16122005 165218 18429 -NCPU=6 -ICPU=3
-DEV TIMEOUT=120 -WARN TIMEOUT=120 -THREADS=1 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV MMAP -LARGEQUERY -NEG_SCORES=0 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21.*
    1: geneseqn1980s:*
    2: geneseqn1990s:*
    3: geneseqn2000s:*
    4: geneseqn2001as:*
    5: geneseqn2001bs:*
    6: geneseqn2002as:*
    7: geneseqn2002bs:*
    8: geneseqn2003as:*
    9: geneseqn2003bs:*
   10: geneseqn2003cs:*
   11: geneseqn2003ds:*
   12: geneseqn2004as:*
   13: geneseqn2004bs:*
   14: geneseqn2005s:*

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1898	100.0	1077	4	AAF29177	Aaf29177 Hydroxyja
2	1898	100.0	1270	3	AAC48839	Aac48839 Arabidops
3	1898	100.0	1273	3	AAC47848	Aac47848 Arabidops
4	1898	100.0	1273	12	ADQ28853	Adq28853 A. thalia
5	1898	100.0	1273	13	ADR21475	Adr21475 Arabidops
6	1553	81.8	1041	4	AAF29178	Aaf29178 Hydroxyja
7	1519	80.0	1023	12	ADQ28854	Adq28854 A. thalia
8	1519	80.0	1023	13	ADR21476	Adr21476 Arabidops
9	1154	60.9	776	3	AAC37672	Aac37672 Arabidops
10	996	52.5	996	6	ABZ13222	Abz13222 Arabidops
11	763.5	40.2	1069	13	ADR59777	Adr59777 Cotton CD
12	757	39.9	980	6	ADA68055	Ada68055 Arabidops
13	757	39.9	981	6	ABZ14244	Abz14244 Arabidops
14	757	39.9	981	6	ADG87892	Adg87892 A. thalia
15	757	39.9	981	6	ADG87829	Adg87829 A. thalia
16	757	39.9	981	8	ABZ42036	Abz42036 Arabidops
17	757	39.9	1043	3	ABZ42382	Abz42382 Arabidops
18	738	38.9	1205	3	AAC37521	Aac37521 Arabidops
19	738	38.9	1212	3	AAC39344	Aac39344 Arabidops
20	680.5	35.9	999	8	ADA69659	Ada69659 Rice gene
21	672	35.4	1440	13	ADR60968	Adr60968 Cotton CD
22	671.5	35.4	271990	10	ADD25213	Add25213 Fertility
23	671.5	35.4	271990	12	ADN61228	Adn61228 Radish nu
24	666.5	35.1	1251	13	ADR60969	Adr60969 Cotton CD
25	666	35.1	1358	13	ADT16337	Adt16337 Plant CDN
26	664.5	35.0	1347	3	AAC37555	Aac37555 Arabidops
27	661.5	34.9	1017	10	AAD54410	Aad54410 Lolium pe
28	654	34.5	1017	3	ABZ68951	Abz68951 Nucleotid
29	654	34.5	1053	10	ABZ68952	Abz68952 Nucleotid
30	653.5	34.4	1053	12	ADN74550	Adn74550 Thale cre
31	653.5	34.4	1053	12	ADN74506	Adn74506 Thale cre
32	653.5	34.4	1281	3	AAC37116	Aac37116 Arabidops
33	653	34.4	1041	10	ABZ68953	Abz68953 Nucleotid
34	653	34.3	1050	11	ACL26221	Acl26221 Rice abio
35	614	32.3	1050	12	ADI45084	Adi45084 Rice isop
36	613.5	32.3	1050	12	ADD25206	Add25206 Fertility
37	613.5	32.3	2067	10	ADN61221	Adn61221 Radish nu
38	601.5	31.7	2067	14	AEB66793	Aeb66793 Rice geno
39	601.5	31.7	1103	8	ADA69350	Ada69350 Rice abio
40	598	31.5	1017	11	ACL26292	Acl26292 Rice abio
41	598	31.5	1017	12	ADI45072	Adi45072 Rice isop
42	593.5	31.3	990	11	ACL26109	Acl26109 Rice abio
43	592	31.2	1403	13	ADX46738	Adx46738 Plant ful
44	592	31.2	1526	13	ADX53055	Adx53055 Plant ful
45	591.5	31.2	1044	11	ACL30321	Acl30321 Rice abio

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
AAF29177
AAF29177 standard; DNA; 1077 BP.

```
ID   AAF29177 standard; DNA; 1077 BP.
XX
AC   AAF29177;
XX
DT   09-APR-2001  (first entry)
XX
DE   Hydroxyjasmonic acid sulfotransferase AtST2a gene.
XX
KW   Hydroxyjasmonic acid sulfotransferase; AtST2a; flowering time; cabbage;
KW   jasmonate; genetically modified plant; lettuce; sugar cane; carrot;
KW   increase vegetative growth; biomass increase; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200102589-A2.
XX
PD   11-JAN-2001.
XX
PF   06-JUL-2000; 2000WO-CA000801.
XX
PR   06-JUL-1999; 99CA-02274873.
XX
PA   (VARI/) VARIN L.
PA   (GIDD/) GIDDA S.
XX
PI   Varin L, Gidda S;
```

```
XX
DR     WPI; 2001-159272/16.
DR     P-PSDB; AAB49722.
XX
PT     Methods for modulating flowering in plants, particularly useful for
PT     plants used in the food-processing industry, involves modifying the
PT     endogenous level of compounds of the jasmonate family.
XX
PS     Claim 38; Fig 7; 50pp; English.
XX
CC     This invention relates to a method for modulating flowering in a plant.
CC     The method comprises modifying the endogenous level of at least one
CC     compound of the jasmonate family in a plant. The methods are used to
CC     produce plants which are genetically modified to flower early or tardily
CC     when compared to a corresponding plant that is not genetically modified,
CC     where the modified plant has an increased (flower early) or lowered
CC     (flower tardily) level of jasmonic acid, or a compound of the jasmonate
CC     family. The method is useful for modulating flowering, particularly for
CC     plants that are used in the food-processing industry and plants with
CC     horticultural value. The method is particularly useful for e.g. delaying
CC     flowering time in crops like lettuce, cabbage, sugar cane or carrots,
CC     which results in increased vegetative growth and biomass. The present
CC     sequence represents the Arabidopsis thaliana AtSrZa gene, which encodes
CC     an hydroxyjasmonic acid sulfotransferase protein, which can be used in
CC     the method of the invention
XX
SQ     Sequence 1077 BP; 294 A; 257 C; 254 G; 272 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:
Score:                   1898.00
Percent Similarity:      100.00%
Best Local Similarity:   100.00%
Query Match:             4

       2.18e-194

              Length:          1077
              Matches:         359
              Conservative:    0
              Mismatches:      0
              Indels:          0
              Gaps:            0

US-10-019-931-3 (1-359) x AAF29177 (1-1077)

Qy         1 MetAlaThrSerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHis 20
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGGCTACCTCAAGCATGAAGAGCATTCCAATGGCAGATCCCAAGTTTTCCATGTGTCAC 60

Qy        21 LysLeuGluLeuLeuLysGlyGlyLysThrArgAspValProLysAlaGluGluAspGlu 40
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 AAGCTCGAGCTCCTAAAGGAAGGCAAAACCCGACGTGTCCCGAAAGCGGAAGAAGATGAA 120

Qy        41 GlyLeuSerCysGluPheGlnGluMetLeuAspSerLeuProLysGlyArgGlyTrpArg 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GGGCTAAGCTGCGAGTTCCAAGAGATGTTGGATTCTCTTCCTAAGGGAGGAGGATGGAGA 180

Qy        61 ThrArgTyrLeuTyrLeuPheGlnGlyIleProPheCysGlnAlaIleAlaIleGlnAlaIle 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 ACTCGTTACCTTTACCTATTCCAAGGGTTTTCGTGCCAAGCAAAGAGATTCAAGCCATC 240

Qy        81 MetSerPheGlnLysHisPheGlnSerLeuGlnAsnAspValValLeuAlaThrIlePro 100
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 ATGTCTCTTTCCAAAAAACATTTCCAAAACGACGTCGTTCTGGCCACCATTACCT 300

Qy       101 LysSerGlyThrThrTrpLeuLysAlaLeuThrPheThrIleLeuAsnArgHisArgPhe 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 AAAATCCGGTACAACCTGGCTAAAAGCTTTAACTTTCACCATCCTTAACCGTCACCGGTTT 360

Qy       121 AspProValAlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAspLeuVal 140
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 GATCCCGGTTGCCTCGAGTACCAACCACCCTCTTTTCACTTCCAACCCTCATGACCTTGTA 420

Qy       141 ProPhePheGluTyrLysAsnGlyAspValProAspLeuSerGlyLeuAla 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 CCTTTCTTCGAGTACAAGCTTTACGCCAACGGAGATGTTCCCGATCTCTCGGGTCTAGCC 480

Qy       161 SerProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluLys 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 AGTCCAAGAACGTTCGCAACCCACTTACCGTTCGGTTCCCTAAAGGAAACGATCGAGAAA 540
```

(continued columns)

```
Db       541 CCCGGCGTGTGAAGGTCCGTTGTACTTGTGTCCCGGAACCCGTTTGACACATTCATCTTGTGG 600

Qy       181 ProGlyValLysValValTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrp 200

Qy       201 HisTyrThrAsnAsnIleLeuSerGluSerValSerProValLeuLeuAsnGlnAlaPhe 220
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 CATTACACCAACAACATCAAATCCGAGTCAGTGAGCCCAGTCCTTGCTCAGACCAAGCTTTT 660

Qy       221 AspLeuTyrCysArgGlyValIleGlyPheGlyProPheTrpGluHisMetLeuGlyTyr 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 GATCTGTATTGCCGGGGAGTGATCGGGTTTGGCCCGTTTTGGGAACACATGTTGGGATAC 720

Qy       241 TrpArgGluSerLeuLysLysArgProGluLysValPhePheLeuArgTyrGluAspLys 260
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 TGGAGAGAGAGCTTGAAGAAGCGACCAGAGAAAGTCTTCTTTTTAAGGTACGAGGATAAA 780

Qy       261 AspAspIleGluThrAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGlu 280
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GACGACATCGAGACAAACTTGAAGAGGCTTGCAACTTTCTTAGAGCTTCCTTTCACCGAA 840

Qy       281 GluGluGluAsnGlyVal 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 GAAGAGGAACAGAAGAGGAGTTGTGAACAAGGCCTATCGCCGAGCTCGTGAGCTTCGGAGAATCTG 900

Qy       301 LysLysLeuGluValAsnLysSerArgIleLeuTyrSerAsnPheGluAsnArgPheLeu 320
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 AAGAAGTTGGAGGTGAACAAGTCAAGAATCCTCTATTCTAACTTCGAGAATCGATTCCTG 960

Qy       321 PheArgLysGlyGluValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArg 340
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 TTTCGGAAAGGAGAAGTGAGTGATTGGGTTAACTATTTGTCACCTTCACAAGTGGAAAGA 1020

Qy       341 LeuSerAlaLeuValAspAspLeuGlyGlySerGlyLeuThrPheArgLeuSer 359
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 TTGTCAGCCTTAGTGGATGACAAGTTAGGTGGAAGTGGTCTCACTTTCAGGTTGAGC 1077
```

```
RESULT 2
AAC48839
ID     AAC48839 standard; DNA; 1270 BP.
XX
AC     AAC48839;
XX
DT     18-OCT-2000 (first entry)
DE     Arabidopsis thaliana DNA fragment SEQ ID NO: 58964.
XX
KW     Hybridisation assay; genetic mapping; gene expression control;
KW     protein identification; signal transduction pathway; metabolic pathway;
KW     promoter; termination sequence; ss.
XX
OS     Arabidopsis thaliana.
XX
PN     EP1033405-A2.
XX
PD     06-SEP-2000.
XX
PF     25-FEB-2000; 2000EP-00301439.
XX
PR     25-FEB-1999;     99US-0121825P.
PR     05-MAR-1999;     99US-0123180P.
PR     09-MAR-1999;     99US-0123548P.
PR     23-MAR-1999;     99US-0125788P.
PR     29-MAR-1999;     99US-0126864P.
PR     01-APR-1999;     99US-0126785P.
PR     01-APR-1999;     99US-0127462P.
PR     06-APR-1999;     99US-0128234P.
PR     08-APR-1999;     99US-0128714P.
PR     16-APR-1999;     99US-0129845P.
PR     19-APR-1999;     99US-0130077P.
PR     21-APR-1999;     99US-0130499P.
PR     23-APR-1999;     99US-0130510P.
PR     23-APR-1999;     99US-0130891P.
```

```
PR 28-APR-1999;   99US-0131449P.
PR 30-APR-1999;   99US-0132048P.
PR 30-APR-1999;   99US-0132407P.
PR 05-APR-1999;   99US-0132484P.
PR 05-MAY-1999;   99US-0132485P.
PR 06-MAY-1999;   99US-0132486P.
PR 06-MAY-1999;   99US-0132487P.
PR 07-MAY-1999;   99US-0132863P.
PR 07-MAY-1999;   99US-0142256P.
PR 11-MAY-1999;   99US-0142256P.
PR 14-MAY-1999;   99US-0142218P.
PR 14-MAY-1999;   99US-0142219P.
PR 14-MAY-1999;   99US-0142221P.
PR 18-MAY-1999;   99US-0134370P.
PR 18-MAY-1999;   99US-0134768P.
PR 19-MAY-1999;   99US-0149941P.
PR 20-MAY-1999;   99US-0135124P.
PR 21-MAY-1999;   99US-0135353P.
PR 24-MAY-1999;   99US-0135629P.
PR 25-MAY-1999;   99US-0136021P.
PR 27-MAY-1999;   99US-0136392P.
PR 28-MAY-1999;   99US-0136782P.
PR 01-JUN-1999;   99US-0137222P.
PR 03-JUN-1999;   99US-0137502P.
PR 04-JUN-1999;   99US-0137528P.
PR 07-JUN-1999;   99US-0137724P.
PR 08-JUN-1999;   99US-0138094P.
PR 10-JUN-1999;   99US-0138540P.
PR 10-JUN-1999;   99US-0138847P.
PR 14-JUN-1999;   99US-0139119P.
PR 16-JUN-1999;   99US-0139452P.
PR 16-JUN-1999;   99US-0139453P.
PR 17-JUN-1999;   99US-0139492P.
PR 18-JUN-1999;   99US-0139454P.
PR 18-JUN-1999;   99US-0139455P.
PR 18-JUN-1999;   99US-0139456P.
PR 18-JUN-1999;   99US-0139457P.
PR 18-JUN-1999;   99US-0139458P.
PR 18-JUN-1999;   99US-0139459P.
PR 18-JUN-1999;   99US-0139460P.
PR 18-JUN-1999;   99US-0139461P.
PR 18-JUN-1999;   99US-0139462P.
PR 18-JUN-1999;   99US-0139463P.
PR 18-JUN-1999;   99US-0139750P.
PR 18-JUN-1999;   99US-0139763P.
PR 18-JUN-1999;   99US-0139817P.
PR 21-JUN-1999;   99US-0139899P.
PR 22-JUN-1999;   99US-0140353P.
PR 23-JUN-1999;   99US-0140354P.
PR 24-JUN-1999;   99US-0140695P.
PR 28-JUN-1999;   99US-0140823P.
PR 29-JUN-1999;   99US-0140991P.
PR 30-JUN-1999;   99US-0141287P.
PR 01-JUL-1999;   99US-0141842P.
PR 01-JUL-1999;   99US-0142154P.
PR 02-JUL-1999;   99US-0142055P.
PR 06-JUL-1999;   99US-0142390P.
PR 08-JUL-1999;   99US-0142803P.
PR 09-JUL-1999;   99US-0142920P.
PR 12-JUL-1999;   99US-0142977P.
PR 13-JUL-1999;   99US-0143542P.
PR 14-JUL-1999;   99US-0143624P.
PR 15-JUL-1999;   99US-0144005P.
PR 16-JUL-1999;   99US-0144085P.
PR 16-JUL-1999;   99US-0144086P.
PR 19-JUL-1999;   99US-0144325P.
PR 19-JUL-1999;   99US-0144331P.
PR 19-JUL-1999;   99US-0144332P.
PR 19-JUL-1999;   99US-0144333P.
PR 19-JUL-1999;   99US-0144334P.
PR 19-JUL-1999;   99US-0144335P.
PR 20-JUL-1999;   99US-0144352P.
PR 20-JUL-1999;   99US-0144632P.
PR 20-JUL-1999;   99US-0144884P.

PR 21-JUL-1999;   99US-0144814P.
PR 21-JUL-1999;   99US-0145086P.
PR 21-JUL-1999;   99US-0145088P.
PR 22-JUL-1999;   99US-0145085P.
PR 22-JUL-1999;   99US-0145089P.
PR 22-JUL-1999;   99US-0145145P.
PR 22-JUL-1999;   99US-0145192P.
PR 23-JUL-1999;   99US-0145145P.
PR 23-JUL-1999;   99US-0145218P.
PR 23-JUL-1999;   99US-0145951P.
PR 26-JUL-1999;   99US-0145224P.
PR 27-JUL-1999;   99US-0145276P.
PR 27-JUL-1999;   99US-0145913P.
PR 28-JUL-1999;   99US-0145918P.
PR 02-AUG-1999;   99US-0145951P.
PR 02-AUG-1999;   99US-0146386P.
PR 02-AUG-1999;   99US-0146388P.
PR 03-AUG-1999;   99US-0146389P.
PR 04-AUG-1999;   99US-0147038P.
PR 04-AUG-1999;   99US-0147204P.
PR 05-AUG-1999;   99US-0147302P.
PR 05-AUG-1999;   99US-0147192P.
PR 06-AUG-1999;   99US-0147260P.
PR 06-AUG-1999;   99US-0147303P.
PR 09-AUG-1999;   99US-0147416P.
PR 09-AUG-1999;   99US-0174493P.
PR 10-AUG-1999;   99US-0147935P.
PR 11-AUG-1999;   99US-0181171P.
PR 12-AUG-1999;   99US-0148319P.
PR 13-AUG-1999;   99US-0148341P.
PR 13-AUG-1999;   99US-0148565P.
PR 16-AUG-1999;   99US-0148684P.
PR 17-AUG-1999;   99US-0149368P.
PR 18-AUG-1999;   99US-0149175P.
PR 18-AUG-1999;   99US-0149929P.
PR 20-AUG-1999;   99US-0149930P.
PR 20-AUG-1999;   99US-0149722P.
PR 23-AUG-1999;   99US-0149902P.
PR 23-AUG-1999;   99US-0149902P.
PR 25-AUG-1999;   99US-0150566P.
PR 26-AUG-1999;   99US-0150884P.
PR 27-AUG-1999;   99US-0151065P.
PR 27-AUG-1999;   99US-0151066P.
PR 30-AUG-1999;   99US-0151080P.
PR 31-AUG-1999;   99US-0151303P.
PR 01-SEP-1999;   99US-0151438P.
PR 07-SEP-1999;   99US-0151930P.
PR 10-SEP-1999;   99US-0153070P.
PR 13-SEP-1999;   99US-0153263P.
PR 15-SEP-1999;   99US-0153758P.
PR 16-SEP-1999;   99US-0154018P.
PR 20-SEP-1999;   99US-0154099P.
PR 22-SEP-1999;   99US-0154779P.
PR 23-SEP-1999;   99US-0155139P.
PR 24-SEP-1999;   99US-0154486P.
PR 28-SEP-1999;   99US-0155659P.
PR 29-SEP-1999;   99US-0154458P.
PR 04-OCT-1999;   99US-0156596P.
PR 05-OCT-1999;   99US-0157177P.
PR 06-OCT-1999;   99US-0157553P.
PR 07-OCT-1999;   99US-0157865P.
PR 08-OCT-1999;   99US-0158029P.
PR 12-OCT-1999;   99US-0158232P.
PR 13-OCT-1999;   99US-0158369P.
PR 13-OCT-1999;   99US-0159293P.
PR 13-OCT-1999;   99US-0159294P.
PR 14-OCT-1999;   99US-0159295P.
PR 14-OCT-1999;   99US-0159299P.
PR 14-OCT-1999;   99US-0159330P.
PR 14-OCT-1999;   99US-0159331P.
PR 14-OCT-1999;   99US-0159637P.
PR 14-OCT-1999;   99US-0159638P.
```

```
PR   18-OCT-1999;  99US-0155584P.
PR   21-OCT-1999;  99US-0160741P.
PR   21-OCT-1999;  99US-0160767P.
PR   21-OCT-1999;  99US-0160768P.
PR   21-OCT-1999;  99US-0160770P.
PR   21-OCT-1999;  99US-0160814P.
PR   21-OCT-1999;  99US-0160815P.
PR   22-OCT-1999;  99US-0160908P.
PR   22-OCT-1999;  99US-0160980P.
PR   22-OCT-1999;  99US-0160981P.
PR   22-OCT-1999;  99US-0160988P.
PR   22-OCT-1999;  99US-0160999P.
PR   25-OCT-1999;  99US-0161404P.
PR   25-OCT-1999;  99US-0161405P.
PR   25-OCT-1999;  99US-0161406P.
PR   26-OCT-1999;  99US-0161359P.
PR   26-OCT-1999;  99US-0161360P.
PR   26-OCT-1999;  99US-0161361P.
PR   28-OCT-1999;  99US-0161920P.
PR   28-OCT-1999;  99US-0161992P.
PR   28-OCT-1999;  99US-0161993P.
PR   29-OCT-1999;  99US-0162142P.
```

Alignment Scores:

```
Pred. No.:
Score:                  1898.00          Length:        1270
Percent Similarity:     100.00%          Matches:        359
Best Local Similarity:  100.00%          Conservative:     0
Query Match:            100.00%          Mismatches:       0
DB:                     3                Indels:           0
                                         Gaps:             0
                        2.76e-194
```

US-10-019-931-3 (1-359) x AAC48839 (1-1270)

```
Qy    1  MetAlaThrSerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHis   20
Db   64  ATGGCTACCTCAAGCATGAAGAGCATTCCAATGGCCATCCCAAGTTTCTCCATGTGTCAC  123

Qy   21  LysLeuGluLeuLeuLysGluGlyLeuTyrThrArgArgAspValProLysAlaGluGlu   40
Db  124  AAGCTCGAGCTCCTTAAAGAAGGGCAAAACTCGACGACGTCCCGAAAGCCGAAGAAGAA   183

Qy   41  GlyLeuSerCysGluPheGlnGlnMetLeuAspSerLeuProLysGluArgGlyTrpArg   60
Db  184  GGGCTAAGCTGCGAGTTCCAAGAGATGTTGGAATTCTCTTCCTAAGGAGAGGAATGGAGA  243

Qy   61  ThrArgTyrLeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysLysGluIleGlnIle   80
Db  244  ACTCGTTACCTTTACCTTATTCCAAGGGTTTTGGTGCCAAGCAAAAGAGATTCAAGCCATC  303

Qy   81  MetSerPheGlnLysHisPheGlnSerLeuGluAsnAspValValLeuAlaThrIlePro  100
Db  304  ATGTCATTTCAAAAACATTTTCAATCCCTCGAAAACGACGGTTGTTCTGGCCACCATACCT  363

Qy  101  LysSerGlyThrThrTrpLeuLysAlaLeuThrPheThrIleLeuAsnArgHisArgPhe  120
Db  364  AAAATCCGGTACAAACCTGGCTAAAAGCTTAACTTCACCATCCTTAACCGTCACCGGTTT   423

Qy  121  AspIleProValAlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAspLeuVal  140
Db  424  GATCCCGTTGCCTCGAGTACCAACACCCTCTTTTCACTTCCAACCCTCATGACCTTGTA   483

Qy  141  ProPhePheGluTyrLysTyrAlaAlaAsnGlyAspValProAspLeuSerGlyLeuAla  160
Db  484  CCTTTCTTCGAGTACAAGCTTACGGCCAACGGAGATGTTCCGGATCTCTCGGGTCTAGCC  543

Qy  161  SerProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluLys  180
Db  544  AGTCCAAGAACGTTCGCCACACATTACCGTTCCTTAAAGGAAACGATCGAGAAA       603

Qy  181  ProGlyValLysValTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrp    200
Db  604  CCCGGTGTGAAGGTCGTACTTGTGCCGGAACCCGTTTGACACATTCATCTCTTCGTGG   663

Qy  201  HisTyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnIleAlaPhe 220
Db  664  CATTACCACCAACACATCAAATCCGAGTCAGTGAGCCCAGTCTTGCTAGACCAAGCTTTT  723

Qy  221  AspLeuTyrCysArgGlyValIleGlyLeuGlyProPheGluHisMetLeuGlyTyr     240
Db  724  GATCTGTATTGCCGGGGAGTGATCGGGTTTGGCCCCGTTTTGGGAACACATGTTGGATAC  783

Qy  241  TrpArgGluSerLeuLysArgProGluLysValPhePheLeuArgTyrGluAspLeuLys  260
Db  784  TGGAGAGAGAGCTTGAAGAGGCCAGAGAAAGTGTTCTTCTTTTTAAGGTACGAGGATCTCAAA 843

Qy  261  AspAspIleGluThrArgAsnLeuLeuLysAlaArgLeuAlaThrPheLeuGluLeuProPheThrGlu 280
Db  844  GACGACATCGAGACCAACCTTAAGGCTCGAAACTTTCTTGAGCTTCCTTTCACCGAA     903

Qy  281  GlyGluGluArgGlyGlyValValGlyAlaIleAlaGluLeuCysSerPheGluAsnLeu  300
Db  904  GAAGAGGAAACGAAAAGGGAGGTTGTGGAAGCCTATCGCGCCGAGCTGTGTAGCTTCGAGAATCTG 963

Qy  301  LysLysGluGluGluArgGlyGlyValValSerAspLysLysSerIleLysLysAsnLysArgPheLeu 320
Db  964  AAGAAGTTGGAGGTGAACAGGTCAAACAAGTGCATCAAGAAGAACTTTGAGAATCGATTCTTG 1023

Qy  321  PheArgLysGlyGluValGlyValSerAsnTyrLeuSerProSerGlnValGluArg     340
Db 1024  TTTCGGAAAGGAGAAGTGAGTGAGTTAACTATTTTGTCACCTTCACACATGGGAAAGA   1083

Qy  341  LeuSerAlaLeuValAspAspAspLysLysGlyLeuGlyLysGlySerGlyLeuThrPheArgLeuSer 359
Db 1084  TTGTCAGCCTTAGTGGATGACAAGAAGGGTGGATCTGGTCTCACTTTCAGGTTGAGC   1140
```

```
PR  07-MAY-1999;  99US-0132863P.
PR  11-MAY-1999;  99US-0134256P.
PR  14-MAY-1999;  99US-0134218P.
PR  14-MAY-1999;  99US-0134219P.
PR  14-MAY-1999;  99US-0134221P.
PR  14-MAY-1999;  99US-0134370P.
PR  18-MAY-1999;  99US-0134768P.
PR  19-MAY-1999;  99US-0134941P.
PR  20-MAY-1999;  99US-0135124P.
PR  21-MAY-1999;  99US-0135353P.
PR  24-MAY-1999;  99US-0135629P.
PR  25-MAY-1999;  99US-0136021P.
PR  27-MAY-1999;  99US-0136392P.
PR  28-MAY-1999;  99US-0136782P.
PR  01-JUN-1999;  99US-0137222P.
PR  03-JUN-1999;  99US-0137528P.
PR  04-JUN-1999;  99US-0137502P.
PR  07-JUN-1999;  99US-0137724P.
PR  08-JUN-1999;  99US-0138094P.
PR  10-JUN-1999;  99US-0138540P.
PR  14-JUN-1999;  99US-0138847P.
PR  16-JUN-1999;  99US-0139119P.
PR  16-JUN-1999;  99US-0139452P.
PR  17-JUN-1999;  99US-0139453P.
PR  18-JUN-1999;  99US-0139492P.
PR  18-JUN-1999;  99US-0139454P.
PR  18-JUN-1999;  99US-0139455P.
PR  18-JUN-1999;  99US-0139456P.
PR  18-JUN-1999;  99US-0139457P.
PR  18-JUN-1999;  99US-0139458P.
PR  18-JUN-1999;  99US-0139459P.
PR  18-JUN-1999;  99US-0139460P.
PR  18-JUN-1999;  99US-0139461P.
PR  18-JUN-1999;  99US-0139462P.
PR  18-JUN-1999;  99US-0139463P.
PR  18-JUN-1999;  99US-0139750P.
PR  21-JUN-1999;  99US-0139763P.
PR  22-JUN-1999;  99US-0139817P.
PR  23-JUN-1999;  99US-0139999P.
PR  23-JUN-1999;  99US-0140353P.
PR  24-JUN-1999;  99US-0140354P.
PR  28-JUN-1999;  99US-0140695P.
PR  29-JUN-1999;  99US-0140823P.
PR  30-JUN-1999;  99US-0140991P.
PR  01-JUL-1999;  99US-0141287P.
PR  01-JUL-1999;  99US-0141842P.
PR  02-JUL-1999;  99US-0142154P.
PR  06-JUL-1999;  99US-0142055P.
PR  08-JUL-1999;  99US-0142390P.
PR  09-JUL-1999;  99US-0142803P.
PR  12-JUL-1999;  99US-0142920P.
PR  13-JUL-1999;  99US-0142977P.
PR  14-JUL-1999;  99US-0143542P.
PR  15-JUL-1999;  99US-0143624P.
PR  16-JUL-1999;  99US-0144005P.
PR  16-JUL-1999;  99US-0144085P.
PR  19-JUL-1999;  99US-0144086P.
PR  19-JUL-1999;  99US-0144325P.
PR  19-JUL-1999;  99US-0144331P.
PR  19-JUL-1999;  99US-0144332P.
PR  19-JUL-1999;  99US-0144333P.
PR  19-JUL-1999;  99US-0144334P.
PR  19-JUL-1999;  99US-0144335P.
PR  20-JUL-1999;  99US-0144352P.
PR  20-JUL-1999;  99US-0144632P.
PR  20-JUL-1999;  99US-0144884P.
PR  21-JUL-1999;  99US-0144814P.
PR  21-JUL-1999;  99US-0144881P.
PR  21-JUL-1999;  99US-0145086P.
PR  21-JUL-1999;  99US-0145088P.
PR  22-JUL-1999;  99US-0145085P.
PR  22-JUL-1999;  99US-0145087P.
PR  22-JUL-1999;  99US-0145089P.
PR  22-JUL-1999;  99US-0145192P.

PR  23-JUL-1999;  99US-0145145P.
PR  23-JUL-1999;  99US-0145218P.
PR  23-JUL-1999;  99US-0145224P.
PR  26-JUL-1999;  99US-0145276P.
PR  27-JUL-1999;  99US-0145913P.
PR  27-JUL-1999;  99US-0145918P.
PR  27-JUL-1999;  99US-0145919P.
PR  28-JUL-1999;  99US-0145951P.
PR  02-AUG-1999;  99US-0146386P.
PR  02-AUG-1999;  99US-0146388P.
PR  02-AUG-1999;  99US-0146389P.
PR  03-AUG-1999;  99US-0147038P.
PR  04-AUG-1999;  99US-0147204P.
PR  04-AUG-1999;  99US-0147302P.
PR  05-AUG-1999;  99US-0147192P.
PR  05-AUG-1999;  99US-0147260P.
PR  06-AUG-1999;  99US-0147303P.
PR  06-AUG-1999;  99US-0147416P.
PR  09-AUG-1999;  99US-0147493P.
PR  09-AUG-1999;  99US-0147935P.
PR  10-AUG-1999;  99US-0148171P.
PR  11-AUG-1999;  99US-0148319P.
PR  12-AUG-1999;  99US-0148341P.
PR  13-AUG-1999;  99US-0148565P.
PR  13-AUG-1999;  99US-0148684P.
PR  16-AUG-1999;  99US-0149368P.
PR  17-AUG-1999;  99US-0149175P.
PR  18-AUG-1999;  99US-0149426P.
PR  20-AUG-1999;  99US-0149722P.
PR  20-AUG-1999;  99US-0149723P.
PR  20-AUG-1999;  99US-0149929P.
PR  23-AUG-1999;  99US-0149902P.
PR  25-AUG-1999;  99US-0149930P.
PR  25-AUG-1999;  99US-0150566P.
PR  26-AUG-1999;  99US-0150884P.
PR  27-AUG-1999;  99US-0151065P.
PR  27-AUG-1999;  99US-0151066P.
PR  30-AUG-1999;  99US-0151080P.
PR  31-AUG-1999;  99US-0151303P.
PR  01-SEP-1999;  99US-0151438P.
PR  07-SEP-1999;  99US-0151930P.
PR  10-SEP-1999;  99US-0152363P.
PR  13-SEP-1999;  99US-0153070P.
PR  15-SEP-1999;  99US-0153758P.
PR  16-SEP-1999;  99US-0154018P.
PR  20-SEP-1999;  99US-0154039P.
PR  22-SEP-1999;  99US-0154779P.
PR  23-SEP-1999;  99US-0155139P.
PR  24-SEP-1999;  99US-0155486P.
PR  28-SEP-1999;  99US-0155599P.
PR  29-SEP-1999;  99US-0155996P.
PR  04-OCT-1999;  99US-0156450P.
PR  05-OCT-1999;  99US-0157117P.
PR  06-OCT-1999;  99US-0157665P.
PR  07-OCT-1999;  99US-0157733P.
PR  08-OCT-1999;  99US-0158029P.
PR  12-OCT-1999;  99US-0158222P.
PR  13-OCT-1999;  99US-0158369P.
PR  13-OCT-1999;  99US-0159293P.
PR  14-OCT-1999;  99US-0159294P.
PR  14-OCT-1999;  99US-0159295P.
PR  14-OCT-1999;  99US-0159299P.
PR  14-OCT-1999;  99US-0159330P.
PR  14-OCT-1999;  99US-0159331P.
PR  18-OCT-1999;  99US-0159377P.
PR  21-OCT-1999;  99US-0159584P.
PR  21-OCT-1999;  99US-0159638P.
PR  21-OCT-1999;  99US-0160741P.
PR  21-OCT-1999;  99US-0160767P.
PR  21-OCT-1999;  99US-0160768P.
PR  21-OCT-1999;  99US-0160770P.
PR  21-OCT-1999;  99US-0160814P.
PR  21-OCT-1999;  99US-0160815P.
```

PR 22-OCT-1999; 99US-0160908P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.:
Score: 2.77e-194
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3

Length: 1273
Matches: 359
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-019-931-3 (1-359) x AAC47848 (1-1273)

```
Qy    1 MetAlaThrSerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHis  20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGGCGACCTCAAGCATGAAGAGCATTCCAATGGCGATCCCAAGTTTCTCCATGTGTCAC  60

Qy   21 LysLeuGluLeuLeuLeuLysGluGlyLysThrArgAspValProLysAlaLeuGluGlu  40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66 AAGCTCGAGCTCCTTAAAGAAGGCAAAAACTCCGACGTCCCGAAAGCCGAAGAAGATGAA 125

Qy   41 GlyLeuSerCysGluPheGlnGluMetLeuAspSerLeuProLysGluArgGlyTrpArg  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  126 GGGCTAAGCTGCGAGTTCCAAGAGATGTTGGATTCTCTTCCTAAGGAGAGAGGATGGAGA 185

Qy   61 ThrArgTyrLeuTyrLeuPheGlnGlyPheThrTrpCysGlnAlaLysGluIleGlnAlaIle  80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  186 ACTCGTTACCTTTACCTTATTCCAAGGGTTTGGTGTGCCAAGCCAAAGAGATTCAAGCCATC 245

Qy   81 MetSerPheGlyHisAlaPheGlnSerLeuGluAsnAspValValLeuAlaThrIlePro 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  246 ATGTCTTTCCAAAAACATTTCCCTCGAAAACACGTGGTTCTCGCACCACCTACT 305

Qy  101 LysSerGlyThrThrTrpLeuGlnThrLeuSerLeuAlaLeuThrPheThrIleLeuAsnArgHisArgPhe 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  306 AAATCCGGTACAAACCTGGCTAAAAGCTTTAACTTTCACCATCACCGGTT 365

Qy  121 AspProValAlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAspLeuVal 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  366 GATCCGGTTGCCTCGAGTACCAACCACACCCTCTTTCACTTCAACCCTCATGACCTTGTA 425

Qy  141 ProPhePheGluThrTyrLysLeuTyrArgLeuAspValProLeuSerGlyTyrLeuAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  426 CCTTTCTTCGAGTACTACAAGCTTTACGCCAACGAGATGTTCCCGATCTCTCGGGTCTTA 485

Qy  161 SerProArgThrPheAlaThrHisLeuProPheGlySerLeuValGluThrIleGluLys 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  486 AGTCCAAGAACGTTCGCAACCCACTTACCGTTCGGTCTTAAGGAAACGATCGAGAAA 545

Qy  181 ProGlyValTyrLysValAlaTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrp 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  546 CCCGGTGTTGAAGGTCGTTACTTGTGCCGGAACCCGTTTGACACATTCATCTCTTCGTGG 605

Qy  201 HisThrTyrThrArgAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaPhe 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  606 CATTACACCTACACAACGAATAACATCAAATCGGAGTCTGTTTCTCCAGTCCTTTTG 665

Qy  221 AspLeuTyrCysArgGlyValIleGlyProTrpGluHisIleMetLeuGlyTyr 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  666 GATCTGTATTGCCGGGGAGTGATCGGGCCTTGGGAACACATAATGTTGGGATAC 725

Qy  241 TrpArgGluSerLeuLysArgProGluLysValPheProheLeuArgGlyTyrGluAspLeuLeuLys 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  726 TGGAGGGAGAGCTTGAAGAGGCCAGAGAAGGTGTTCCCTTTTTTAAGGTACGAGGATCTCAAA 785
```

```
Db  786 TGGAGGAGAGCGTTGAAGAGGCCAGAGAAGGTGTTCCTTTTTTAAGGTACGAGGATCTCAAA 845

Qy  261 AspAspIleGluThrArgAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGlu 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  846 GACGACATCGAGACGACCAACTTGAAGAGAGGCTTGCAAACTTTCTTTAGAGCTTTCACCGAA 905

Qy  281 GluGluGluArgLysGlyValValLysLysAlaIleAlaGluLeuCysSerPheGluAsnLeu 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  906 GAAGAGGAAAGGAAAGGAGTTGTGAAGAAAGCTATCGCGGAGCTTTGTAGCTTCGAGAATCTG 965

Qy  301 LysLysGluValAlaSerLeuAsnLysSerIleIleLysGluAsnProGluGluAsnArgPheLeu 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  966 AAGAAGTTGGAGGTGACAAGTCAAGAACAAGTCCAATCAAGAAACTTTGAGAATCGATTCTTG 1025

Qy  321 PheArgLysGlyGlyValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArg 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1026 TTTCGGAAAGGAGGAGTGAGTGATTGGGTTAACTATTTGTCACCTTCACAAGTTGGAAAGA 1085

Qy  341 LeuSerAlaLeuValLeuAspAspAspProLysLeuGlyLysSerGlyTyrLeuThrPheArgLeuSer 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1086 TTGTCAGCCCTTAGTGGATGACGACAAGTTAGGTGGATCTGGTCTCACTTTCAGGTTGAGC 1142
```

RESULT 4
ADQ28853
ID   ADQ28853 standard; cDNA; 1273 BP.
XX
AC   ADQ28853;
XX
DT   23-SEP-2004 (first entry)
XX
DE   A. thaliana sulfotransferase family protein AtST2a cDNA SEQ ID NO:1.
XX
KW   tuber size; tuber; plant; hydroxy-jasmonate sulfotransferase; enzyme;
KW   jasmonate; Arabidopsis thaliana; sulfotransferase family protein; AtST2a;
KW   chromosome 5; gene; ss.
XX
OS   Arabidopsis thaliana.
XX
FH
FT   Key   Location/Qualifiers
FT   CDS   66..1145
FT   /*tag= a
FT   /product= "sulfotransferase family protein AtST2a"
XX
PN   WO200405700-A2.
XX
PD   08-JUL-2004.
XX
PF   23-DEC-2003; 2003WO-CA002009.
XX
PR   23-DEC-2002; 2002CA-02415062.
XX
PA   (FLOR-) FLORISYS INC.
XX
PI   Varin L;
XX
PT   Modulating tuber size in tuber producing plant such as Solanum tuberosum,
PT   involves modifying in plant endogenous level of tuber modulating compound
PT   such as jasmonate by modifying activity of hydroxy-jasmonate
PT   sulfotransferase.
XX
DR   WPI; 2004-517683/49.
DR   P-PSDB; ADQ28855.
DR   GENBANK; NM_120783.
XX
PS   Claim 4; SEQ ID NO 1; 32pp; English.
XX
CC   The present invention describes a method (M1) for modulating the tuber
CC   size in a tuber producing plant. (M1) involves modifying in the plant the
CC   endogenous level of at least a tuber modulating compound by modifying the
CC   activity of a hydroxy-jasmonate sulfotransferase, the tuber modulating
CC   compound being a jasmonate. Also described; (1) a plant cell
CC   transformation vector (I) capable of facilitating transfer and expression

CC of an exogenous nucleic acid into an isolated cell and/or facilitating
CC integration of an exogenous nucleic acid into genome of the cell,
CC comprising at least one promoter sequence, one enhancer sequence and one
CC exogenous nucleic acid sequence, the promoter being a constitutive
CC expression promoter or an inducible promoter, the exogenous nucleic acid
CC being chosen from nucleic acid sequence or its antisense, having at least
CC 50% homology with the Arabidopsis thaliana sulfotransferase family
CC protein AtST2a or the AtST2b sequence having 1023 nucleotides (S1, ADQ28853) or the
CC AtST2b sequence having 1023 nucleotides (S2, ADQ28854), and a nucleic
CC acid encoding an amino acid sequence having at least 50% homology with
CC the AtST2a sequence of 359 amino acids (S3, ADQ28855) or the AtST2b
CC sequence of 340 amino acids (S4, ADQ28856); and (2) a genetically
CC modified tuber producing plant obtained by (M1) in which the endogenous
CC level of the 11- or 12-hydroxyjasmonic acid sulfotransferase is higher
CC (when sense exogenous nucleic acid is used) or lower (when antisense
CC exogenous nucleic acid is used), compared to the endogenous level of the
CC 11- or 12- hydroxyjasmonic acid sulfotransferase in a non genetically
CC modified tuber producing plant. (M1) and (I) can be used for modulating
CC tuber size in a tuber producing plant such as Solanum tuberosum, Solanum
CC demissum, Ipomea batatas, Dioscorea batatas, Helianthus tuberosus and
CC Astragalus complanatus. The present sequence encodes the Arabidopsis
CC thaliana sulfotransferase family protein AtST2a, which is used in the
CC exemplification of the present invention. The AtST2a gene is located on
CC chromosome 5.
XX
SQ Sequence 1273 BP; 369 A; 296 C; 280 G; 328 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:
Score: 1898.00
Percent Similarity: 100.00%
Percent Identity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB:
DB: 12

US-10-019-931-3 (1-359) x ADQ28853 (1-1273)

Sequence 1273 BP; 369 A; 296 C; 280 G; 328 T; 0 U; 0 Other;

Length: 1273
Matches: 359
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

Qy   1 MetAlaThrSerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHis 20
Db   66 ATGGCTACCTCAAGCAATGAAGAGCAATTCCAATGGCGATTCCCAAGTTTCTCCATGTGTCAC 125

Qy   21 LysLeuGluLeuLysLysGluGlyLysThrArgArgValProLysAlaGluGluAspGlu 40
Db   126 AAGCTCGAGCTCCTTAAAGAGGGCAAAACTCGGCACGTCCCCGAAGCCGAAGAAGATGAA 185

Qy   41 GlyLeuLeuSerCysGluPheGlnGlyMetLeuAspSerLeuProLeuGluArgGlyTrpArg 60
Db   186 GGGCTTAAGCTCGAGTTCCAAGGAATGTTGGATTCTCTTCCTAAGGAGAGGGGATGGAGA 245

Qy   61 ThrArgTyrLeuTyrLeuPheGlnGlyPheThrCysGlnAlaLysGluIleGlnAlaIle 80
Db   246 ACTCGTTACCTTTACCTATTCCAAGGGTTTACTTGTCAGGCAAAAGAGATTCAAGCCATC 305

Qy   81 MetSerPheGlnLysHisPheGlnSerLeuGluAsnAspValValLeuAlaThrIlePro 100
Db   306 ATGTCTTTCCAAAAACATTTCCAATCCCTCGAAAACGACGTCGTTCTCGCCACCATACCT 365

Qy   101 LysSerGlyThrThrTrpLeuLysGluAlaLeuThrPheMetThrIleLeuAsnArgHisArgPhe 120
Db   366 AAATCCGGTACAACCTGGCTAAAAGCTTTAACTTTCATGACCATCCTGAACCGGTTT 425

Qy   121 AspProValAlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAspLeuVal 140
Db   426 GATCCGGTTGCCTCGAGTACCAACCACCCTCTTTTCACTTCCAACCCTCATGACCTTGTA 485

Qy   141 ProPhePheGluThrTyrLysLeuTyrAlaAsnGlyAspValProAspLeuSerGlyLeuAla 160
Db   486 CCTTTCTTCGAGTACAAGCTTTACGCCAACGGAGATGTTCCCGATCTCTCGGGTCTAGCC 545

Qy   161 SerProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluLys 180
Db   546 AGTCCAAGAACGTTCGCAACCCACTTACCGTTCGGTTCCCTAAAGGAAACCATCGAAAA 605

PT hydroxyjasmonate by increasing in plant, level of in-vivo sulfonation of
PT hydroxyjasmonates or decreasing level of synthesis of 11-and/or 12-
PT hydroxyjasmonate.

XX

PS Claim 6; SEQ ID NO 1; 29pp; English.

XX

CC The invention relates to a method for producing (M1) a male sterile
CC plant, involving decreasing the level of 11-and/or 12-hydroxyjasmonate by
CC increasing in the plant the level of in-vivo sulfonation of
CC hydroxyjasmonate or decreasing the level of synthesis of 11-and/or 12-
CC hydroxyjasmonate. Further disclosed is a plant cell transformation vector
CC (I) capable of facilitating transfer and expression of an exogenous
CC nucleic acid into an isolated cell and/or facilitating integration of an
CC exogenous nucleic acid into genome of the cell. In (M1), the level of in-
CC vivo sulfonation of hydroxyjasmonate is increased by increasing in the
CC plant the endogenous activity of a hydroxyjasmonate sulfotransferase. The
CC increasing of the level of in-vivo sulfonation of hydroxyjasmonates or
CC the decreasing of the level of synthesis of 11-and/or 12-
CC hydroxyjasmonate is achieved by a process chosen from genetic
CC modification of the plant, chemical mutagenesis of the plant and
CC selection of natural mutants, preferably genetic modification. The
CC endogenous activity of the sulfotransferase is increased by stimulating
CC the expression of one or more gene chosen from Arabidopsis thaliana
CC sulfotransferase gene (AtST2a) and AtST2b. The method of the invention is
CC useful for producing a male sterile plant. It is useful for restoring
CC normal anther development in a genetically modified male sterile plant.
CC The current sequence represents the Arabidopsis sulfotransferase (AtST2a)
CC encoding gene.

XX

SQ Sequence 1273 BP; 369 A; 296 C; 280 G; 328 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:   2.77e-194
Score: 1898.00
Percent Similarity: 100.00%
Percent Identity: 100.00%
Best Local Similarity: 100.00%
Query Match: 13

DB:

US-10-019-931-3 (1-359) x ADR21475 (1-1273)

Qy   1 MetAlaThrSerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHis 20
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGGCTACCTCAAGCATGAAGAGCATTCCAATGGCAATCCCAAGTTTCTCCAATGTGTCAC 125

Qy   21 LysLeuGluLeuLeuLysGluGlyLysThrArgAspValProLysAlaGluGluAspGlu 40
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   66 AAGCTCGAGCTCCTTAAAGAAGGCAAAAACTCGGCAGTGCCCGAAAGCCGAAGAGATGAA 185

Qy   41 GlyLeuSerGlyPheGlnGluMetLeuAspSerLeuProLysGlnArgGlyTyrArg 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 126 GGGCTAAGCTGCGAGTTCCAAGATGTTGGATCTTCTCTTCCTAAGCAGAGAGGGATGGAGA 245

Qy   61 ThrArgTyrLeuPheGlnGlyPheTrpTyrSerGlnAlaAlaAlaGluIleGlnAlaIle 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 186 ACTCGTTACCTTTACCTATTCCAAGGGTTTTGGTCGCCAAGCCAAGAGATTCAAGCCATC 305

Qy   81 MetSerPheGlnLysHisPheGlnSerLeuGluAsnAspValValLeuAlaThrIlePro 100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 246 ATGTCTTTCCAAAAACATTTCCAATCCCTCGAAAACGACGTCGTTCTGCCCACCATACCT 365

Qy 101 LysSerGlyThrThrTrpLeuLysAlaLeuThrPheThrIleLeuAsnArgHisArgPhe 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 306 AAGTCCGGTACCACTTGGCTAAAGGCTTTAACTTTCACCATCCTTAACCGTCACCGGTTT 425

Qy 121 AspProValAlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAspLeuVal 140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 366 AAAATCCGGTACCAACCTGGCTAAAAGCTTTAACTTTCACCATCCTTAACCGTCACCGGTTT 485

Qy 141 ProPhePheGluThrTyrAlaAlaAsnGlyLysValProAspLeuSerGlyLeuAla 160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 426 GATCCGGTTGCCTCGAGTACCAACCACCTCTTTTCACTTCCAAACCCTCATGACCTTGTA 545

Qy 161 SerProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluGlyLys 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 486 CCTTTCTTCGAGTACGCCAACGGAAGAGATGTTCCGAACCTTAGCGGTCTAGCC 545

Qy 161 ProGlyValLysValValThrTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrp 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 546 AGTCCAAGAACGTTCGCAACCCACTTACCGTTCCCTAAAGGAAAACGATCGAGAAA 605

Qy 181 ProGlyValLysValValThrTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrp 200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 546 AGTCCAAGAACGTTCGCAACCCACTTACCGTTCCCTAAAGGAAAACGATCGAGAAA 605

Qy 201 HisTyrThrArgAsnIleLysSerGluSerValSerProValLeuLeuAspThrGlnAlaPhe 220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 606 CCCGGGTGTGAACGGTCGTGTACTTGTGCCGGGACCGTTTGACACATTCATCTCTTCGTGG 665

Qy 221 AspLeuTyrCysArgGlyValAlaIleGlyProTrpPheGluHisIleMetGluGlyTyr 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 666 CATTACACCAGAAACATCAAAAGTCAGAGTGCCGGTCAGTCCTGCTAGACCAAGCTTTT 725

Qy 241 TrpTyrArgGluSerLeuLysProGluLysValPhePheLeuArgTyrGluAspLeuLys 260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 726 GATCTGTATTGCCGGGGAGTGATCGGGCCGTTTGGGAAACACATGTTGGAGGGATAC 785

Qy 261 AspAspIleGluThrAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGlu 280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 786 TGGAGAGAGAGCTTGAAGCCGGAGAAAGTCTTCTTTTTAAGGTACGAGGATCTCAAA 845

Qy 281 GluGluLeuArgLysGlyValValLysTyrCysSerPheGluAsnLeu 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 846 GACGACATCGAGACCAACTTGAAGAGGCTTGCAACTTTCTTAGAGCTTCCTTTCACCGAA 905

Qy 301 LysLysLeuGluValAspAsnLysSerAsnLysSerIleLeuLysAsnPheGluLeu 320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 906 GAAGAGGAACGAAAGGGAGTTGTGAAGTACTGCTCGTTCGAGAATCTG 965

Qy 321 PheArgLysGlyGluValSerAspTrpValAsnLysSerIleLeuSerProSerGlnValGluArg 340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 966 AAGAAGTTGGGAGGTGAACAAGTCAAACAACTTTGAGAATCGATTCTTG 1025

Qy 341 LeuSerAlaLeuValAspAspAspGlyLeuGlySerGlyLeuThrPheArgGluSer 359
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1026 TTTCGGAAGGAGGAAGTGAGTGATTGGGTTAACTATTTGTCACTTCACAAGTGGAAAGA 1085

Qy   --
Db 1086 TTGTCAGCCTTAGTGGATGACAAGTTAGGTGGATTAGGATCTGGTCTCACTTTCAGGTTGAGC 1142

RESULT 6

AAF29178
ID   AAF29178 standard; DNA, 1041 BP.
XX
AC   AAF29178;
XX
DT   09-APR-2001 (first entry)
XX
DE   Hydroxyjasmonic acid sulfotransferase AtST2b gene.
XX
KW   Hydroxyjasmonic acid sulfotransferase; AtST2b; flowering time; cabbage;
KW   jasmonate; genetically modified plant; lettuce; sugar cane; carrot;
KW   increase vegetative growth; biomass increase; ds.
XX
OS   Arabidopsis thaliana.
XX
XX
PN   WO200102589-A2.
XX
PD   11-JAN-2001.
XX
PF   06-JUL-2000; 2000WO-CA000801.
XX
PR   06-JUL-1999;   99CA-02274873.
XX
PA   (VARI/) VARIN L.
PA   (GIDD/) GIDDA S.
XX
PI   Varin L, Gidda S;
XX
PR   WPI; 2001-159272/16.
DR   P-PSDB; AAB49723.
XX
PT   Methods for modulating flowering in plants, particularly useful for

```
PT   plants used in the food-processing industry, involves modifying the
PT   endogenous level of compounds of the jasmonate family.
XX
PS   Claim 38; Fig 9; 50pp; English.
XX
CC   This invention relates to a method for modulating flowering in a plant.
CC   The method comprises modifying the endogenous level of at least one
CC   compound of the jasmonate family in a plant. The methods are used to
CC   produce plants which are genetically modified to flower early or tardily
CC   when compared to a corresponding plant that is not genetically modified,
CC   where the modified plant has an increased (flower early) or lowered
CC   (flower tardily) level of jasmonic acid, or a compound of the jasmonate
CC   family. The method is useful for modulating flowering, particularly for
CC   plants that are used in the food-processing industry and plants with
CC   horticultural value. The method is particularly useful for e.g. delaying
CC   flowering time in crops like lettuce, cabbage, sugar cane or carrots,
CC   which results in increased vegetative growth and biomass. The present
CC   sequence represents the Arabidopsis thaliana AtST2b gene, which encodes
CC   an hydroxyjasmonic acid sulfotransferase protein, which can be used in
CC   the method of the invention
XX
SQ   Sequence 1041 BP; 287 A; 241 C; 242 G; 271 T; 0 U; 0 Other;
```

US-10-019-931-3 (1-359) x AAP29178 (1-1041)

```
Qy     11  MetAlaIleProSerPheSerMetCysHisHisLysLeuGluLeuLeuLysGluGlyLysThr  30
           |||||||||||||||||||||||||||||||||||||||||||||||| ::||||||||||
Db      1  ATGGCGATCCCAAGTTTCTCCATGTGTCACCACAAGCTCGAGCTCCTTAAGGAAGGCAAAAGC  60

Qy     31  ArgAspValProLysAlaGluAspGluAspGlyLeuSerCysPheGlnGluMetLeu  50
           :::  |||||||||||||||||||||||||||||||||
Db     61  --------GAAGGCCAAGAAGAAGAGGGGCTTAAGCTACAGAGTTCCAAGAGATGTTG  108

Qy     51  AspSerLeuProLysGluLysGlyTyrTrpArgThrArgTyrLeuTyrLeuPheGlnGlyPhe  70
           |||||||||||||::||||||||||||||||||||||||||||||||||||||||||||||||
Db    109  GACTCTCTTCCTAAGGAGAAGGGACCGGAGAAATCGTTACCTTTACTTATTCCAAGGGGTTT  168

Qy     71  TrpCysGlnAlaLysGlyIleGlnIleAlaIleMetSerPheGlyInLysHisPheGlnSerLeu  90
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169  CGGTGCCCAAGCTAAGGAGGATTCAAGCTATCACGTCTTTCACAAAAACATTTTCAGTCCCTT  228

Qy     91  GluAsnAspValValLeuValAlaThrIleProLysSerGlyThrThrTrpLeuLysAlaLeu  110
           :::|||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db    229  CCAGACGAGGTTGTCCTGGCCACCATAACCTAAAATCCTGGGTTAAAAGCTTTA  288

Qy    111  ThrPheThrIleLeuAsnArgHisArgPheAspProVal---AlaSerSerThrAsnHis  129
           ||||||||||::||||||||||||||||||||||||    ::||||||::::||
Db    289  ACTTTCACCATCCTTACCGGTTGATCCGGTTCCATCAAGTTCCGACCAC  348

Qy    130  ProLeuPheThrSerAsnProHisAspLeuValProPheGluTyrIleLysLeuTyrAla  149
           ||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db    349  CCTCTTCTCACATCCAACCCTCTACCTTTCTTCGAGTACAAGCTTTAGGCC  408

Qy    150  AsnGlyAspValProAspLeuSerGlyTyrLeuAlaSerProArgThrPheAlaThrHisLeu  169
           ||||:::||||||||||||||||||||||::|||||||||||||||||::::|||
Db    409  AACGGAAATGTTCCGGATCTCTCGGGTTCAGCCAGACAATTCCAGCACAGTA  468

Qy    170  ProPheGlySerLeuLysGluThrIleGluLysValAlaValValTyrLeuCys  189
           |||||||||||||||:::::|||||||::|||||||||||||||||||
Db    469  CCGTTCGGTTCGCTTAAGGAAATCCAGTGTAGGTTGTGTACCTGTGC  528

Qy    190  ArgAsnProPheAspThrPheIleSerSerTrpHisTyrThrArgAsnLeuLysSerGlu  209
           |||||||||||||||||||||||||||||:::||||||||||||||||||||||||||
Db    529  CGGAACCCGTTGACACATTCATCTCCATGTGGCATTACATCAACAACATCACTTCCGAG  588

Qy    210  SerValSerProValLeuLeuAsnGlnAlaPheAspLeuTyrCysArgGly---ValIle  228
           ||||||||||||||||||||||||||||||||||||||||||||||| ::||||
Db    589  TCAGTGAGCCCAGTCTTGCTAGGACGAAGCCTTTGATCTATATTGCCCGGGGATTACTGATC  648

Qy    229  GlyPheGlyProTrpTrpGluHisMetLeuGlyTyrTrpArgGluSerLeuLysArgPro  248
           |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db    649  GGATTTGGGCCGTGGTGGGAACACATGTTGGGATACTGGAGAGAGCTTGAAGAGGCCA  708

Qy    249  GluLysValPhePheArgLeuArgThrTyrGluAspAspIleGluThrArgAsnLeuLys  268
           ||||||||||::::||||||||||||||||||||||||||||||||||||||||||||||
Db    709  GAGAAAGTCTTTATTTTAAAGTACGAGGATGACATCGAGACAAGGAACTTGAAG  768

Qy    269  ArgLeuAlaThrPheLeuGluLeuProPheThrGluGluGluLeuAsnPheValVal  288
           |||||||::||||:||||||||||||||||||||::||||||||||||||::||||||||
Db    769  AAGCTAGCAACGTTCTTAGGACTTCCTTTCACCGAAGAAGAGCTGAACTTTGTGTG  828

Qy    289  LysAlaIleAlaGluLeuLeuCysSerPheGluAsnLeuLeuLysLysLysLeuAspAsnLysSer  308
           |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db    829  AAAGCTATCGCGTGAGCTTCTGAGAATCTGAGAAAAATCGAAGAAGTTTGAGAACAAGTCA  888

Qy    309  AsnLysSerIleLysAsnPheGluAsnArgPheLeuProAspGlyGluValSerAsp  328
           :::|||:::||||||||||||||||||||||||||||||||||||||||
Db    889  AGCAAATTGATCCAGAACTACAGAACCGGTTCTTGTTAGGAGAAGGAGAAGTCAT  948

Qy    329  TrpValAsnTyrLeuSerProSerGlnValGluArgLeuSerAlaLeuValAspAspLys  348
           :::|||:::||||||||||||||||::|||||||||||||||||||||||||||||
Db    949  TTGGTTAACTATTTGTCGCCATCACAAGTGGAAAGATTGTCAGCCCTTAGTCGAGACAAG  1008

Qy    349  LeuGlyGlySerGlyLeuThrPheArgLeuSer  359
           |||||||||||||||||||||||||||||||||
Db   1009  TTAGCTGGATCTGGGTCTCCACTTTCAGATTGAGT  1041
```

RESULT 7
ADQ28854

```
PT   sulfotransferase.
XX
PS   Claim 5; SEQ ID NO 2; 32pp; English.
XX
CC   The present invention describes a method (M1) for modulating the tuber
CC   size in a tuber producing plant. (M1) involves modifying in the plant the
CC   endogenous level of at least a tuber modulating compound by modifying the
CC   activity of a hydroxy-jasmonate sulfotransferase, the tuber modulating
CC   compound being a jasmonate. Also described: (1) a plant cell
CC   transformation vector (I) capable of facilitating transfer and expression
CC   of an exogenous nucleic acid into an isolated cell and/or facilitating
CC   integration of an exogenous nucleic acid into genome of the cell,
CC   comprising at least one promoter sequence, one enhancer sequence and one
CC   exogenous nucleic acid sequence, the promoter being a constitutive
CC   expression promoter or an inducible promoter, the exogenous nucleic acid
CC   being chosen from nucleic acid sequence or its antisense, having at least
CC   50% homology with the Arabidopsis thaliana sulfotransferase family
CC   protein AtST2a sequence having 1023 nucleotides (S1, ADQ28853) or the
CC   AtST2b sequence having 1023 nucleotides (S2, ADQ28854), and a nucleic
CC   acid encoding an amino acid sequence having at least 50% homology with
CC   the AtST2a sequence of 359 amino acids (S3, ADQ28855) or the AtST2b
CC   sequence of 340 amino acids (S4, ADQ28856); and (2) a genetically
CC   modified tuber producing plant, obtained by (M1), in which the endogenous
CC   level of the 11- or 12-hydroxyjasmonic acid sulfotransferase is higher
CC   (when sense exogenous nucleic acid is used) or lower (when antisense
CC   exogenous nucleic acid is used), compared to the endogenous level of the
CC   11- or 12- hydroxyjasmonic acid sulfotransferase in a non genetically
CC   modified tuber producing plant. (M1) can be used for modulating the
CC   tuber size in a tuber producing plant such as Solanum tuberosum, Solanum
CC   demissum, Ipomea batatas, Dioscorea batatas, Helianthus tuberosus and
CC   Astragalus complanatus. The present sequence encodes the Arabidopsis
CC   thaliana sulfotransferase family protein AtST2b, which is used in the
CC   exemplification of the present invention. The AtST2b gene is located on
CC   chromosome 5.
XX
SQ   Sequence 1023 BP; 285 A; 234 C; 238 G; 266 T; 0 U; 0 Other;


US-10-019-931-3 (1-359) x ADQ28854 (1-1023)

Alignment Scores:
Pred. No.:                       1.47e-153
Score:                           1519.00
Percent Similarity:              92.15%
Best Local Similarity:           84.88%
Query Match:                     80.03%
DB:                              12
```

```
Qy    18 MetCysHisLysLeuGluLeuLeuLysGluGlyLysThrArgAspValProLysAlaGlu 37
          ||| ::|| ||||||
Db     1 ATGTGTCACAAGCCCGAGCTCCTTAAGGAAGGCAAAAGC------GAAGGCCAA 48

Qy    38 GluAspGlyLysGlyLeuSerCysGluLeuPheGlnMetLeuAspSerLeuProLysGluArg 57
          ||| |||||:||||||||||||||||||||||||||
Db    49 GAAGAGGAGGGCTAAGCTACGAGTTCCAAGAGATGTTGGACTCTCTTCCTAAGGAGAGA 108

Qy    58 GlyTrpArgThrArgTyrLeuTyrLeuPheGlnGlyThrTrpCysGlnAlaLysLysGlyIle 77
          |||||||||||||||||||||||||||||||||||||||||
Db   109 GGACCGGAGAAATCGTTACCTTTACTTATTCCAAGGGTTTCGGTGCCCAAGCTAAGGAGATT 168

Qy    78 GlnAlaIleMetSerSerPheGlnLysHisSerLeuGluAsnAspValValLeuAla 97
          |||||||||||||||||||||||||||||||||||||||
Db   169 CAAGCTATCACGTCTTTCCAAAAACATTTTCAGTCCCTTCGAGAACGAGGTTGTCCTGGCC 228

Qy    98 ThrIleProLysSerGlyThrThrTrpLeuLysGluThrThrIleLeuAsnArg 117
          |||||||:|||||||||||||||||||||||||||||
Db   229 ACCATACCTAAAAACTGGGACAACCTGGTTAAAGCTTTAAACTTTCACCATCCTTACCCGGT 288

Qy   118 HisArgPheAspProVal---AlaSerSerThrAsnHisSerProLeuPheThrSerAsnPro 136
          |||||||||||||     ::||||||||||||
Db   289 CATCGGTTTGATCCGGTTTCCTCATCAAGTTCCGACCACCCTCTTCTCACATCCAACCCT 348

Qy   137 HisAspLeuValProPhePheGluTyrLysLeuTyrAlaAsnGlyAspValAspProAspLeu 156
          |||||||||||||||||||||||||||||||||||||||::||||||
Db   349 CACGACCTCGTACCTTTCTTCGAGTACAAGCTTTACCCGCAACAGGAAATGTTCCGGATCTC 408

Qy   157 SerGlyLeuAlaSerProArgThrPheAlaThrHisLeuProHisGlySerLeuLysGlu 176
          ||||||||||||||||::||||||||||||||||||
Db   409 TCGGGTCCTTAGCCAGTCCAAGAACATTCGCACCCACCTACCCGTTCGGTGCCCTTAAGGAT 468

Qy   177 ThrIleGlyLysProGlyValLysValTyrLeuCysArgAsnProPheAspThrPhe 196
          ||||||||||||||||||||||||||||||||||::||||
Db   469 TCGGTCGAGAATCCAGTGTGAAGGTTGTGTACCTGTGCCGGAACCCGTTTGACACATTC 528

Qy   197 IleSerSerTrpHisTyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeu 216
          ||||||||||||||||||||||||||||||||||||||||||
Db   529 ATCTCCAGTTGGCATTACACATCAACAACATCACTTCCGAGTCAGTGAGCGCAGTCTTGCTA 588

Qy   217 AspGlnAlaPheAsnLeuTyrCysArgGly---ValIleGlyProPhePheTrpGlu 235
          |||||||||||||||||||||     ::|||||||||||||
Db   589 GACCAAGCTTTTGAATCTATATTGCCGGGGGATTACTGACGTTCGGTGCCCGTTTTGGGAA 648

Qy   236 HisMetLeuGlyTyrTrpArgGluSerLeuLysProLysGlyArg 255
          |||||||||||||||::||||||||||||||||||
Db   649 CACATGTTGGGATACTGGAGAGAGAGCTTGAAGACGAAGGCCCAGAGAAAAGTCTTATTTTAAAG 708

Qy   256 TyrGluAspLeuLysAspAspSerIleGluThrArgAsnLeuLeuLysArgLeuAlaThrGluPheLeuGlu 275
          |||||||||||||||||:::|||||||||||||||::::||||
Db   709 TACGAGGATCTCAAAGAAGACAGCATCGAGACCCGCAACCTGCTGAAGAGACTTGCAACAGAGTTTCTTGAGGA 768

Qy   276 LeuProPheThrGluGluGluGlyGlyValValAlaLeuLysGlyValValAlaIleIleAlaGluLeuCys 295
          |||||||||::|||||||||||||||::|||||||||||||||
Db   769 CTTCCTTTCACCGAAGAAGAAGAGGGAGGCGTTGTGGCATTGAAAGGTGTTGTAGCTATCGCTGATCTGTGT 828

Qy   296 SerPheGluAsnLeuLysLeuGluValValAsnLysSerArgAsnLysSerIleIleGluLysAsnPhe 315
          |||||||||||||||||::|||||||||||||||||||||
Db   829 AGCTTTGAGAATCTGAAGCTTGAGGTTGTGAACAAGTCAAGAAATAAGTCGATCATAGAGAAGAACTTT 888

Qy   316 GluAsnArgPheLeuProPheArgLysGlyLysValGlyValSerAspTrpValAsnTyrLeuSerPro 335
          ||| :|||||||||||||||||||||||||||||::|||||||
Db   889 GAGAACCGGTTCTTGCCTTTCAGGAAAGGAAAAGTCGGTGATTGGGTTAACTATTTGTCGCCA 948

Qy   336 SerGlnValGluArgLeuSerAlaLeuValAlaAspArgLysLeuGlyGlySerGlyLeuThr 355
          |||||||||||||||||||||||||||||||||||||||
Db   949 TCACAAGTCGAAAGATTGTCAGCCTTAGTGGCTGATAGAAAGCTTGGTGGTTCTGGTCTCACT 1008

Qy   356 PheArgLeuSer 359
          ||||||||||||
Db  1009 TTCAGATTGAGT 1020
```

```
RESULT 8
ADR21476
ID   ADR21476 standard; DNA; 1023 BP.
XX
AC   ADR21476;
XX
DT   21-OCT-2004 (first entry)
XX
DE   Arabidopsis sulfotransferase (AtST2b) encoding gene, seq id 2.
XX
KW   Male sterile plant; 11-hydroxyjasmonate; 12-hydroxyjasmonate;
KW   sulfonation; sulfotransferase gene; AtST2a; AtST2b; plant; gene; ds.
XX
OS   Arabidopsis thaliana.
XX
FH   Key              Location/Qualifiers
FT   CDS              1..1023
FT                    /*tag= a
FT                    /product= "AtST2b protein"
XX
PN   WO2004062350-A2.
XX
PD   29-JUL-2004.
XX
PF   13-JAN-2004; 2004WO-CA000006.
XX
PR   13-JAN-2003; 2003CA-02414487.
```

```
XX
PA   (FLOR-) FLORISYS INC.
PA   (PFLA-) INST PFLANZENBIOCHEMIE.
XX
PI   Maucher H, Miersch O, Wasternack C, Varin L;
XX
XX
DR   WPI; 2004-571390/55.
DR   P-PSDB; ADR21478.
XX
PT   Producing male sterile plant, involves decreasing level of 11-and/or 12-
PT   hydroxyjasmonate by increasing in plant, level of in-vivo sulfonation of
PT   hydroxyjasmonates or decreasing level of synthesis of 11-and/or 12-
PT   hydroxyjasmonate.
XX
PS   Claim 6; SEQ ID NO 2; 29pp; English.
XX

CC   The invention relates to a method for producing (M1) a male sterile
CC   plant, involving decreasing the level of 11-and/or 12-hydroxyjasmonate by
CC   increasing in the plant the level of in-vivo sulfonation of 11-and/or 12-
CC   hydroxyjasmonates or decreasing the level of synthesis of 11-and/or 12-
CC   hydroxyjasmonate. Further disclosed is a plant cell transformation vector
CC   (I) capable of facilitating transfer and expression of an exogenous
CC   nucleic acid into an isolated cell and/or facilitating integration of an
CC   exogenous nucleic acid into genome of the cell. In (M1), the level of in-
CC   vivo sulfonation of hydroxyjasmonates, is increased by increasing in the
CC   plant the endogenous activity of a hydroxyjasmonate sulfotransferase. The
CC   increasing of the level of in-vivo sulfonation of hydroxyjasmonate or
CC   the decreasing of the level of synthesis of 11-and/or 12-
CC   hydroxyjasmonate is achieved by a process chosen from genetic
CC   modification of the plant, chemical mutagenesis of the plant and
CC   selection of natural mutants, preferably genetic modification. The
CC   endogenous activity of the sulfotransferase is increased by stimulating
CC   the expression of one or more gene chosen from Arabidopsis thaliana
CC   sulfotransferase gene (AtST2a) and AtST2b. The method of the invention is
CC   useful for producing a male sterile plant. It is useful for restoring
CC   normal anther development in a genetically modified male sterile plant.
CC   The current sequence represents the Arabidopsis sulfotransferase (AtST2b)
CC   encoding gene.
XX
SQ   Sequence 1023 BP; 285 A; 234 C; 238 G; 266 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:               1.47e-153
Score:                   1519.00        Length:        1023
Percent Similarity:      92.15%         Matches:        292
Percent Identity:        84.88%         Conservative:    25
Best Local Similarity:   80.03%         Mismatches:      21
Query Match:             13             Indel:            6
DB:                                     Gaps:             3

US-10-019-931-3 (1-359) x ADR21476 (1-1023)

Qy   18  MetCysHisLysLeuGluLeuLeuLysGluGlyLysThrArgAspValProLysAlaGlu  37
             |||||||||||||||||||||||||||||||||||||:||
Db    1  ATGTGTCACAAGCCCGAGCTTCCTTAAGGAGGGCAAAAGC---------GAAGGCCAA  48

Qy   38  GluAspGluGlyLeuSerCysGluPheGlnGluMetLeuAspSerLeuProLysGluArg  57
Db   49  GAAGAGGAGGGTCTAAGCTGCGAGTTCCAAGAGATGTTGGACTCTCTTCCTAAGGAGAGA 108

Qy   58  GlyTrpArgThrArgTyrLeuTyrLeuHisPheThrProCysGlnAlaLysGluIle     77
Db  109  GGACGGAGAAATCGTTACCTTTACTTTATTCCAAGGGTTTCGGTCGCCAAGTTAAGGAGATT 168

Qy   78  GlnAlaIleMetSerPheGlnLysHisPheGlnSerLeuGluAsnAspValValLeuAla  97
Db  169  CAAGCTATCACGTCTTTCCAAAAACATTTTCAGTCCCTTGGAGAACGATGTTGTCCTCGCC 228

Qy   98  ThrIleProLysSerGlyTyrThrPheLeuLysSerAlaLeuThrPheThrIleLeuAsnArg 117
Db  229  ACCATACCTAAATCTGGCACAACCTTCACTTTAACTTTCACCATCCTTACCGT         288

Qy  118  HisArgPheAspProVal---AlaSerSerThrAsnHisProLeuPheThrSerAsnPro 136
```

```
Db  289  CATCGGTTTGATCCGGGTTTCCTCATCAAGTTCCGACCACCCTTCTCTCACATCCAACCCT 348

Qy  137  HisAspLeuValProPhePheGluTyrLysLeuTyrLysAlaAsnGlyAspValProLeu 156
Db  349  CACGACACCTCGTTACCTTTCTTCGAGTACAAGCTTTACAAGGCAAATGTTCCCGGATCTC 408

Qy  157  SerGlyLeuValAlaSerProArgThrPheHisAlaThrHisLeuProPheGlySerLeuLysGlu 176
Db  409  TCGGGTCTTGTAGCCAGTCCCAGGACATTCCACGCAACCCACTTACCGTTCGGGTCTTTAAGGAT 468

Qy  177  ThrIleGlyLysProGlyValValLysValTyrLeuCysArgAsnProPheAspThrPhe 196
Db  469  TCGGTCGAGAATCCCAGTGTGAAGGTTTGTACCTGTCGGGAAACCCGTTTGACACATTC 528

Qy  197  IleSerSerTrpHisTyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeu 216
Db  529  ATCTCCATGTGGCATTACACTAACAACATCACTTCCGAGTCAGTGAGCCCAGTTCTTGCTA 588

Qy  217  AspGlnAlaPheAspLeuThrTyrCysArgGly---ValIleGlyPheGlyProPheTrpGlu 235
Db  589  GACGAACGTTTGATCTTRATATGCCGGGGATTTACATGGGCCCGTTTTGGGAA         648

Qy  236  HisMetLeuGlyTyrTrpArgGluSerLeuLeuLysGlyArgProGluLysValPhePheLeuArg 255
Db  649  CACATGTTGGGATACTGGAGAGAGAGCTTGAAGGGCCGACCGGAGAAAGTCTTATTTTTAAAG 708

Qy  256  TyrGluAspLeuLysLysAspProValThrAsnLeuLysArgLeuGluAlaThrPheLeuGlu 275
Db  709  TACGAGGATCTCAAAAGAAGAACATCGAGACCAACTTGAAGAGACCAGTAACCAGAGAAAGTTCTTAGGA 768

Qy  276  LeuProPheThrGluGluGluGluArgLysGlyValValAlaLysAlaIleAlaGluLeuCys 295
Db  769  CTTCCTTTCACCGAAGAAGAGGAACAAAAGGGAGTTGTGAAAGCTATCGCTGAACTCTGT 828

Qy  296  SerPheGluAsnLeuLysSerAsnLysSerIleLysLysAsnPhe                 315
Db  829  AGCTTTGAGAATCTGAAAAGTAACAAAAGCATCAAGAAGAACTTT                 888

Qy  316  GluAsnArgPheLeuPheArgLysGlyGluValSerAspTrpValAsnTyrLeuSerPro 335
Db  889  GAGAACCGGTTCTTGTTTTGGAAAAGGAGAGGTGAGTCAATTTGGTTAACTTATTTGTCGGCA 948

Qy  336  SerGlnValGlyArgLeuSerAlaLeuValAlaAspAspGlyLysGlySerGlyLeuThr 355
Db  949  TCACAAGTGGAAAGATTGTCAGCCTTAGTGGATGACAAGGTTAGCTGGATCTGGTCTCACT 1008

Qy  356  PheArgLeuSer 359
Db 1009  TTCAGATTGAGT 1020
```

```
RESULT 9
AAC37672
ID   AAC37672 standard; DNA; 1154 BP.
XX
AC   AAC37672;
XX
DT   17-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 18246.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999; 99US-01218255P.
```

```
PR  05-MAR-1999;  99US-0123180P.
PR  09-MAR-1999;  99US-0125548P.
PR  23-MAR-1999;  99US-0125788P.
PR  25-MAR-1999;  99US-0126264P.
PR  29-MAR-1999;  99US-0126785P.
PR  01-APR-1999;  99US-0127462P.
PR  06-APR-1999;  99US-0128234P.
PR  08-APR-1999;  99US-0128714P.
PR  16-APR-1999;  99US-0129845P.
PR  19-APR-1999;  99US-0130077P.
PR  21-APR-1999;  99US-0130449P.
PR  23-APR-1999;  99US-0130510P.
PR  28-APR-1999;  99US-0130891P.
PR  30-APR-1999;  99US-0131449P.
PR  30-APR-1999;  99US-0132048P.
PR  04-MAY-1999;  99US-0132407P.
PR  05-MAY-1999;  99US-0132484P.
PR  06-MAY-1999;  99US-0132485P.
PR  06-MAY-1999;  99US-0132486P.
PR  06-MAY-1999;  99US-0132487P.
PR  07-MAY-1999;  99US-0132863P.
PR  11-MAY-1999;  99US-0134256P.
PR  14-MAY-1999;  99US-0134218P.
PR  14-MAY-1999;  99US-0134219P.
PR  14-MAY-1999;  99US-0134221P.
PR  14-MAY-1999;  99US-0134370P.
PR  18-MAY-1999;  99US-0134768P.
PR  19-MAY-1999;  99US-0134941P.
PR  20-MAY-1999;  99US-0135124P.
PR  21-MAY-1999;  99US-0135353P.
PR  24-MAY-1999;  99US-0135629P.
PR  25-MAY-1999;  99US-0136021P.
PR  27-MAY-1999;  99US-0136392P.
PR  28-MAY-1999;  99US-0137222P.
PR  01-JUN-1999;  99US-0137528P.
PR  03-JUN-1999;  99US-0137502P.
PR  04-JUN-1999;  99US-0137724P.
PR  07-JUN-1999;  99US-0138094P.
PR  08-JUN-1999;  99US-0138540P.
PR  10-JUN-1999;  99US-0138847P.
PR  14-JUN-1999;  99US-0139119P.
PR  16-JUN-1999;  99US-0139452P.
PR  16-JUN-1999;  99US-0139453P.
PR  17-JUN-1999;  99US-0139492P.
PR  18-JUN-1999;  99US-0139454P.
PR  18-JUN-1999;  99US-0139455P.
PR  18-JUN-1999;  99US-0139456P.
PR  18-JUN-1999;  99US-0139457P.
PR  18-JUN-1999;  99US-0139458P.
PR  18-JUN-1999;  99US-0139459P.
PR  18-JUN-1999;  99US-0139460P.
PR  18-JUN-1999;  99US-0139461P.
PR  18-JUN-1999;  99US-0139462P.
PR  18-JUN-1999;  99US-0139463P.
PR  18-JUN-1999;  99US-0139750P.
PR  21-JUN-1999;  99US-0139763P.
PR  22-JUN-1999;  99US-0139817P.
PR  23-JUN-1999;  99US-0139899P.
PR  23-JUN-1999;  99US-0140353P.
PR  23-JUN-1999;  99US-0140354P.
PR  24-JUN-1999;  99US-0140695P.
PR  28-JUN-1999;  99US-0140823P.
PR  29-JUN-1999;  99US-0140991P.
PR  30-JUN-1999;  99US-0141287P.
PR  01-JUL-1999;  99US-0141842P.
PR  01-JUL-1999;  99US-0141854P.
PR  02-JUL-1999;  99US-0142154P.
PR  06-JUL-1999;  99US-0142055P.
PR  08-JUL-1999;  99US-0142390P.
PR  08-JUL-1999;  99US-0142803P.
PR  09-JUL-1999;  99US-0142920P.
PR  12-JUL-1999;  99US-0142977P.
PR  13-JUL-1999;  99US-0143542P.

PR  14-JUL-1999;  99US-0143624P.
PR  15-JUL-1999;  99US-0144005P.
PR  16-JUL-1999;  99US-0144085P.
PR  16-JUL-1999;  99US-0144086P.
PR  19-JUL-1999;  99US-0144325P.
PR  19-JUL-1999;  99US-0144331P.
PR  19-JUL-1999;  99US-0144332P.
PR  19-JUL-1999;  99US-0144333P.
PR  19-JUL-1999;  99US-0144334P.
PR  19-JUL-1999;  99US-0144335P.
PR  20-JUL-1999;  99US-0144352P.
PR  20-JUL-1999;  99US-0144632P.
PR  20-JUL-1999;  99US-0144804P.
PR  21-JUL-1999;  99US-0144814P.
PR  21-JUL-1999;  99US-0144806P.
PR  22-JUL-1999;  99US-0145085P.
PR  22-JUL-1999;  99US-0145087P.
PR  22-JUL-1999;  99US-0145089P.
PR  22-JUL-1999;  99US-0145092P.
PR  23-JUL-1999;  99US-0145145P.
PR  23-JUL-1999;  99US-0145218P.
PR  26-JUL-1999;  99US-0145224P.
PR  27-JUL-1999;  99US-0145276P.
PR  27-JUL-1999;  99US-0145913P.
PR  28-JUL-1999;  99US-0145918P.
PR  02-AUG-1999;  99US-0145919P.
PR  02-AUG-1999;  99US-0145951P.
PR  03-AUG-1999;  99US-0146388P.
PR  04-AUG-1999;  99US-0146389P.
PR  05-AUG-1999;  99US-0147303P.
PR  06-AUG-1999;  99US-0147416P.
PR  09-AUG-1999;  99US-0147493P.
PR  09-AUG-1999;  99US-0147935P.
PR  10-AUG-1999;  99US-0148171P.
PR  11-AUG-1999;  99US-0148319P.
PR  12-AUG-1999;  99US-0148341P.
PR  13-AUG-1999;  99US-0148565P.
PR  16-AUG-1999;  99US-0148684P.
PR  17-AUG-1999;  99US-0149175P.
PR  18-AUG-1999;  99US-0149426P.
PR  20-AUG-1999;  99US-0149722P.
PR  20-AUG-1999;  99US-0149723P.
PR  23-AUG-1999;  99US-0149902P.
PR  23-AUG-1999;  99US-0149930P.
PR  25-AUG-1999;  99US-0150566P.
PR  26-AUG-1999;  99US-0150884P.
PR  27-AUG-1999;  99US-0151065P.
PR  27-AUG-1999;  99US-0151066P.
PR  30-AUG-1999;  99US-0151080P.
PR  31-AUG-1999;  99US-0151303P.
PR  01-SEP-1999;  99US-0151438P.
PR  07-SEP-1999;  99US-0151930P.
PR  10-SEP-1999;  99US-0152363P.
PR  13-SEP-1999;  99US-0153070P.
PR  15-SEP-1999;  99US-0153758P.
PR  16-SEP-1999;  99US-0154018P.
PR  20-SEP-1999;  99US-0154039P.
PR  22-SEP-1999;  99US-0154779P.
PR  23-SEP-1999;  99US-0155139P.
PR  24-SEP-1999;  99US-0155486P.
PR  28-SEP-1999;  99US-0155599P.
PR  29-SEP-1999;  99US-0155658P.
PR  29-SEP-1999;  99US-0155996P.
PR  04-OCT-1999;  99US-0157117P.
```

PR   05-OCT-1999;   99US-0157753P.
PR   06-OCT-1999;   99US-0157786P.
PR   07-OCT-1999;   99US-0158029P.
PR   08-OCT-1999;   99US-0158232P.
PR   12-OCT-1999;   99US-0158836P.
PR   13-OCT-1999;   99US-0159293P.
PR   13-OCT-1999;   99US-0159294P.
PR   13-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159931P.
PR   14-OCT-1999;   99US-0159677P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   26-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161992P.
PR   29-OCT-1999;   99US-0162142P.

Alignment Scores:

Pred. No.:
Score: 776.00
Percent Similarity: 61.49%
Best Local Similarity: 41.09%
Query Match: 40.89%
DB: 3

Length: 1154
Matches: 143
Conservative: 71
Mismatches: 122
Indels: 12
Gaps: 4

US-10-019-931-3 (1-359) x AAC37672 (1-1154)

Qy   17   SerMetCysHisLys----------------------LeuGluLeuLeuLysGluGlyLys   29
           |||.||||:||.|| ::||:||||||||.||||||||
Db   57   TCTTTGTTGTTTCAAATCGTGGGGAACAAATTATCACAATGGAGGCTTCTAAAGAA------   110

Qy   30   ThrArgAspValProLysAlaGlyGluGlyLyaGlyLeuSerCysGluPheGlnGluMet   49
           ||:::::||||.||||||||.|||||.:||.|||||||||:||||.|:|.|
Db   111 GCTCATCACCTTCCAAACTACATGAAAGACGACAACGTTAGTCAACAGAAACCAAGAACTTG   170

Qy   50   LeuAspSerLeuProLysGluArgGlyTrpArgThrArgTyrLeuPheGlnGly   69
           |||||||||.|||.||:|||||||||||:|||||:
Db   171 ATCACTTCTCTACCTTCAGCACAAAGATTTCATGGGTTATGGTTCTTCTTACAACTACAAAGGT   230

Qy   70   PheTrpCysGlnAlaLysGluIleGlnAlaIleMetSerPheGlyHisPheGlnSer   89
           :||||||||:||||:||:||||||
Db   231 TGTTGGTGTACTATCCAAACACACTCCAAGCCGTTCTTGACGTCCAAAAACACTTCAGCCA   290

Qy   90   LeuGluAsnAspValValLeuAlaThrIleProLysGlyThrThrTrpLeuLysAla   109
           ||.::||||
Db   291 CGAGATACTGATAATAATCCTGGCTTCTTTGCCCAAAGGTGGAACACCACTTGGCTCAAATCC   350

Qy   110 LeuThrPheThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThrAsnHis   129
           |||||||:|:|||.|||
Db   351 CTAATTTTCCGCTGTTGTACATAGAGAGAAAAGTACCGC-----GGAACCCTTCAAACACAT   404

Qy   130 ProLeuPheThrSerAsnProHisAspLeuValProPheGlyThrLysLeuTyrLysAla   149
           |||.|||.||||
Db   405 CCTTTTGCTTCTTACAAAACCCTACAAAACCCTCATGACCTTGTCCCATTTCTTGAGGTTATACGCT   464

Qy   150 AsnGlyAspValProAspLeuSerGlyLeuAlaSerProAlaThrArgThrPheAlaThrHisLeu   169
           ||||||:|||||:|||||:|||

RESULT 10
AB213222
ID   AB213222 standard; DNA; 996 BP.
XX
AC   AB213222;
XX
DT   21-JAN-2003 (first entry)
XX
DE   Arabidopsis thaliana stress regulated gene SEQ ID NO 1027.
XX
KW   Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
OS   Arabidopsis thaliana.
XX
PN   WO200216655-A2.
XX
PD   28-FEB-2002.
XX
PF   24-AUG-2001; 2001WO-US026685.
XX
PR   24-AUG-2000; 2000US-02278666P.
PR   26-JAN-2001; 2001US-0264647P.
PR   22-JUN-2001; 2001US-0300111P.
XX
PA   (SCRI) SCRIPPS RES INST.
PA   (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI   Harper JF, Kreps J, Wang X, Zhu T;
XX
DR   WPI; 2002-304127/34.

Db   465 AATAGCCAAAATTCCGGATCTCGGAAAGTATTCTTCTCCTATAGATCTTTTCTACACACATG   524
           ||| ::|||||| |||::|:||||
Qy   170 ProPheGlySerLeuLysGluThrIleGluLysProGlyValLysValValTyrLeuCys   189
           :::||||||| |||:||||:||::|||
Db   525 CACTTACAAGCATTGCGTGAACGCCACCACAAA---GCTTGCAAAACCGTAATGTGTGT   581

Qy   190 ArgAsnProPheAspThrPheIleSerSerTrpHisTyrThrAsnAsnIleLysSerGlu   209
           ||| ||||||:|||||||||:|:||||| |||||||||
Db   582 AGAGGTATCAAAAGATACGTTTGCTCTCCTGGTCATTATAGAACATGTTGCATCGCACC   641

Qy   210 SerValSerProValLeuLeuAspGlyAlaPheAspLeuTyrCysArgGlyValIleGly   229
           ||| :||||||||||||.|||:|||||||||:||||:
Db   642 AAGATGGATCAAGCCACTTTGATCTCATGTTTATGATGCGTTATTGTAGAGGAGTTCTTTA   701

Qy   230 PheGlyProPheTrpGluHisMetLeuGlyTyrTyrArgGluSerLeuLeuLysArgProGlu   249
           :::|||||||||||||| |||||||||:::|:||||||:||
Db   702 TACGGACCTTATTGGGAACATGTATTGGAGTATTATAGAGAAAGCAACAGAG   761

Qy   270 LeuAlaThrPheLeuGluGluLeuProPheThrGluGluLeuGluLysGlyLysValValLys   289
           |||:||||||||||||||||||| |:|:|||||||:||:|:||||
Db   762 AATGTTCTTTTCATGAAGTACCGAAGAGATAATTGAGGAGCCTCGTGTTCCAAGTCAAGAGA   821

Qy   290 AlaIleAlaGluLeuCysSerPheGluAsnLeuLysGluLysLeuValLeuAsnLysSerAsn   309
           :::||||||:|||||||||||:|:||||||:|||||||||||:||||||||||||
Db   822 CTCGCCGGAGTTCTTGGAAATGTCCATTCACCAGGAAGAAGAGAAAGTGGATCGGTGGAG   881

Qy   310 LysSerIleLysAsnPheLeuAsnArgLysGlyGlyGluValSerAspTrp   329
           :::||||||||||||| |||||||||||:|:||||||:||||||
Db   882 GAGATCTTGAAGTTGTTGTAGTTTACGAAATTTAAGCAAATTTGAGGTTAATAAGAATGG   941

Qy   330 ValAsnTyrLeuSerProSerGlnValGluArgLeuSerAlaLeuValAlaSerAspLysLeu   349
           |||:::||||||||||||||| |||:|:||||||::|||||
Db   942 ACAACGAGAAATTGGTGTGTAGGATTCTCAGGTGTTGTTCTTTAGGAGGGTGGAAGTTGGTGATGG   1001

Qy   350 GlyGlySerGlyLeuThrPheArg   357
           ||||||||:|||| |:||
Db   1002 AAGAATCATCTTACGCCACAAAATGGCAAAAACCTTTGATGAGATTATTGACTATAGACTA   1061

Db   1062 GGAGACTCCGGTTTGATATTTCA   1085

XX
PT Identifying a stress condition to which a plant cell has been exposed and
producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 1027; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 996 BP; 306 A; 199 C; 213 G; 278 T; 0 U; 0 Other;

US-10-019-931-3 (1-359) x AB213222 (1-996)

Alignment Scores:
Pred. No.:
Score: 7.04e-73   Length: 996
Score: 770.50   Matches: 140
Percent Similarity: 62.50%   Conservative: 70
Best Local Similarity: 41.67%   Mismatches: 121
Query Match: 40.60%   Indels: 5
DB: Gaps: 3

Qy   22 LeuGluLeuLeuLysGlyGlnGluGlyLysThrArgAspValProLysGlyAlaProLeuGlyGly 41
           :::|||
Db   1 .....ATGGAGGCTTCTAAAGAA------GCTCATCACCTTCAAACTACATCGAAAACGAACAAC 54

Qy   42 LeuSerCysGlyProPheGlnGluMetLeuAspSerLeuProLysGlyLeuArgGlyTyrArgThr 61
           :::|:||
Db   55 GTTAGTTCAAGAAGCCAAGAAACCAAGAACTTGAACCACTTCCTTACCTTACAGAAGAATTCATGGGT 114

Qy   62 ArgTyrTyrLeuPheGlyCysGlnAlaLysGlyLeuIleGlnAlaIleMet 81
           |||||||:::
Db 115 TATGGGTCTTCTACAACTACAAAAGGTTGTTGGTACTATATCCAAACACACTCCAAGCCGTTCTT 174

Qy   82 SerPheGlnLysHisPheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLys 101
           :::|||||||
Db 175 GACGTCCAAAAACACTTCAAGCCACGAGATACTGTATATAAATCCTCGCTTTCTTTGCCCAAA 234

Qy 102 SerGlyThrThrThrLeuThrTrpLeuLysSerAlaLeuGluThrPhePheThrIleLeuLeuAsnArgHisArgPheAsp 121
           :::||||||
Db 235 GGTGGAACCACTTGGCTCAAAATCCCTAATTTTCGCGTTGTACATAGGAGAAATGGTACCGC 294

Qy 122 ProValAlaSerSerThrAsnHisProLeuPheThrSerAsnHisAspLeuSerLeuValPro 141
           :::
Db 295 ------GGAACCCTCAAACACATCCTTTGCTCTTACAAAACCCTCATGACCTTGTCCCA 348

Qy 142 PhePheGluTyrTyrArgLysLeuTyrAlaAsnGlyAspValProAspLeuSerGlyLeuAlaSer 161
           :::|||||||
Db 349 TTTCTTCGAGGTTGAGTTATACGCTAATAGCCAAAATTCCGGATCTTCGCAAAGTATTCTTCT 408

Qy 162 ProArgThrPheAlaThrHisLeuProPheGlySerLeuGlyThrIleGlyLysPro 181
           ||| :::
Db 409 CCTATGATCTTTTCTACACAATGGCGTTCAAGCACCACCACAAA-- 465

Qy 182 GlyValValLysValThrTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrpHis 201
           :::||||||
Db 466 GCTTGCAAACCGTATATGTGTGGAGGTATCAAAGATACGTTTGTCTCCGGCCAT 525

Qy 202 TyrThrArgAsnLysGluSerGluSerValSerProValLeuLeuValAspGlnAlaPheAsp 221
           ||| :::
Db 526 TATAGAAACATGTGCATCCGCACCAAGATGGATCAAGCCACTTTGAGCTCAGTTTGAT 585

Qy 222 LeuTyrCysSerArgGlyValIleGlyPheGlyProGluHisMetLeuGlyTyrTrp 241
           |||||||:::|||
Db 586 GCTTATGTGTAGAGGAGGTTCTCTTATATGGGACTTATTGGGGGACATGTGATTGACTTATTGG 645

Qy 242 ArgGluSerArgLeuLysArgProGluLysGluValPhePheLeuArgTyrGluAspLeuLysAsp 261
           :::|||
Db 646 AAAGGGGAGCTTGGAAGCAAAAGAAGAGAGAATTGTTCTTTTCATGAGATAGCAGAAGAGATAAATTGAG 705

Qy 262 AspIleGluThrAsnLeuLysArgLeuAlaThrPheLeuGluGluLeuGluGluThrGluGlu 281
           :::|||||||
Db 706 GAGCCTCGTGTTCAAGTCAAGAAGACTCCCGGAGTTCTTGGAATGTCCAATTCACCCAAGGAA 765

Qy 282 GluGluArgLysGlyValValLysValAlaIleLeuAlaGluLeuCysSerPheGluAsnLeuLys 301
           |||||||:::
Db 766 GAAGAAGAAAAGTGGATCGGTGGAAGGAGATCTTGAAGTTGTGTTACGAAAATTTAAGC 825

Qy 302 LysLeuGluValAsnLysSerAsnLysSerIleIleLysAsnPheGluAsnArgPheLeuPhe 321
           ||| :::
Db 826 AATTTGGAGGTTAATAAGAATGGGACACGAGAATTGGTAGATTCTCAGGTGTTCTTTT 885

Qy 322 ArgLysGlyGluValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArgLeu 341
           |||||||:::|||
Db 886 AGGAAGGGTGAAGTTGGGTGGAATTGGAAGAAATCATCTTACGCCACAAAATGGCGAAAACCTTT 945

Qy 342 SerAlaLeuValAspAspProLysLeuGlySerGlyLeuThrPheArg 357
           :::::::|||||||
Db 946 GATGAGGATTATTGACTATAGACTCCGGTTTGATAATTTCAA 993

RESULT 11
ADR59777
ID   ADR59777 standard; cDNA; 1069 BP.
XX
AC   ADR59777;
XX
DT   02-DEC-2004 (first entry)
XX
DE   Cotton cDNA sequence, SEQ ID 558.
XX
KW   Cotton; ssi; plant; cold tolerance; growth rate; cell cycle pathway;
KW   drought tolerance; plant disease resistance; galactomannan; lignin;
KW   plant growth regulator; heat tolerance; herbicide tolerance;
KW   homologous recombination; extreme osmotic condition tolerance;
KW   pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
KW   stress resistance.
XX
OS   Gossypium hirsutum.
XX
PN   US20040181830-A1.
XX
PD   16-SEP-2004.
XX
PF   29-JAN-2004; 2004US-00767795.
XX
PR   07-MAY-2001; 2001US-00849529.
PR   12-DEC-2001; 2001US-00021323.
XX
PA   (KOVA/) KOVALIC D K.
PA   (ZHOU/) ZHOU Y.
PA   (CAOY/) CAO Y.
XX
PI   Kovalic DK, Zhou Y, Cao Y;
XX
DR   WPI; 2004-667718/65.
XX
PT   New recombinant nucleic acid molecules and polypeptides from Gossypium
PT   hirsutum, useful for producing plants with improved biological
PT   characteristics (e.g. improved plant cold or drought tolerance).
XX
PS   Claim 1; SEQ ID NO 558; 14pp; English.
XX
CC   The invention relates to a recombinant polynucleotide comprising any of
CC   the 58798 Cotton plant cDNA sequences mentioned in the specification.
CC   Also a recombinant polypeptide comprising any of the 58798 amino acid
CC   sequences mentioned in the specification and producing a plant having an
CC   improved property. Producing a plant having an improved property

CC   comprises transforming a plant with a recombinant construct comprising a
CC   promoter region functional in a plant cell operably joined to a
CC   polynucleotide comprising a coding sequence for a polypeptide associated
CC   with the property, and growing the transformed plant. The polypeptide is
CC   useful for improving plant cold tolerance, manipulating growth rate in
CC   plant cells by modification of the cell cycle pathway, improving plant
CC   drought tolerance, providing increased resistance to plant disease,
CC   producing galactomannan (or lignin or plant growth regulators), improving
CC   plant heat tolerance, improving plant tolerance to herbicides, increasing
CC   the rate of homologous recombination in plants, improving plant tolerance
CC   to extreme osmotic conditions or to pathogens or pests, improving yield
CC   by modification of photosynthesis, modifying seed oil or protein yield
CC   and/or content, improving yield by modification of carbohydrate, nitrogen
CC   or phosphorus use and/or uptake, or improving yield by providing improved
CC   plant growth and development under at least one stress condition. The
CC   polynucleotide and polypeptide may also be used in recombinant DNA
CC   constructs, in physical arrays of molecules, as plant breeding markers,
CC   or in computer-based storage and analysis systems. The present sequence
CC   is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20040181830. However only 6585
CC   polynucleotide sequences were available, the remaining 52213
CC   polynucleotides and all 58798 protein sequences were not present.
XX

SQ   Sequence 1069 BP; 318 A; 252 C; 210 G; 289 T; 0 U; 0 Other;

```
Alignment Scores:
Pred. No.:
Score:                  763.50            Length:         1069
Percent Similarity:     64.17%            Matches:         139
Percent Identity:       43.30%            Conservative:     67
Best Local Similarity:  40.23%            Mismatches:      110
Query Match:            13                Indels:            5
                                          Gaps:              2
```

US-10-019-931-3 (1-359) x ADR59777 (1-1069)

```
Qy    12 AlaIleProSerPheSerMetCysHisLysGluLeuGluLeuLeuGlyGluGlyLysThrArg    31
            |:::||||||:||         ::            ||||
Db   119 TCAATTACCAGCTGAAAATGATTGCCAT-----------ACTCAAGCTGAAAGAATTCAGTTTCT   169

Qy    32 AspValProLysAlaGluGluAspGlyGlyLeuSerCysGluPheGlnGluMetLeuAsp    51
            |||||||||||:|           :::            |||||||||:::
Db   170 GATGTTCCCAAAATACCTGCTAGAATTGGGTCTCACCCCAGAATGCAAGCAGCTTATATCC   229

Qy    52 SerLeuProLysGluArgGlyTrpAsgThrArgTyrLeuPheLeuPheGluGlyPheTrp    71
            |||||||||:|||||||:|||:|||:||
Db   230 TCACTGCCAGCGCGAGAAAGGTTGGGTTGCCAATCTCCTCCATCAATATCAGGGGTTTGG   289

Qy    72 CysGlnAlaLysGluIleGlnAlaIleIleMetSerSerPheGlnLysHisPheGlnSerLeuGlu    91
            |||:|||||:||||||:||
Db   290 CACACCACTCCGCCAGTTACAGCAGTGCTTACTTGCCAAAACCACTTTCAGCTCAGGAG   349

Qy    92 AsnAspValValLeuAlaThrIleProLysSerGlyTyrThrThrLysProLeuLeuLeuThr   111
            |||:||::|||||||         ||
Db   350 ACAGATATCCCTCCTTGTTACCACTCCAAAATCAGGAGACAACAGTTAAAGGCTATTGTT   409

Qy   112 PheThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThrArgAsnHisProLeu   131
            |||||||:|||:||||||||||||       |||||
Db   410 TTTGCCCTTGATGAACCCAGTCAAGTACCCC------AACACTGATAACAATCACCCTTG   463

Qy   132 PheThrSerGlyArgAsnProHisAspLeuValProPhePheGluThrTyrLysLeuTyrAlaAsnGly   151
            ::|          ||||||||||||||||||||||||||||||||| :::
Db   464 CTCTCAGAGAACCCTCACAATTCTTGTTCCATTTTGGAGTATGGGGCTTTACATTGATAGT   523

Qy   152 AspValProAspLeuSerGlyLeuAlaSerProArgThrPheAlaThrHisLeuProPhe   171
            |||||||:|||::|            |||||||||||||||
Db   524 CAGGTTCCTAACTTACACCACTTTACATCTCCACGGCTATCGGAACTCATTACCCCTT   583

Qy   172 GlySerLeuLysGluThrIleGlyLysProGlyValLysValTyrLeuCysArgAsn   191
            |||::|||||||||||||||:||:||||:||:||||||||||:||:
Db   584 GTTTCAATTGCCAGAATCGGCAAAGAACTCATCTTGCAAGCTGTTTATTTAGTAGGAAC   643

Qy   192 ProPheAspThrPheIleSerSerTrpHisTyrThrAsnAsnIleLysSerGluSerVal   211
            ||||||||||:|::||||:::                    ::
Db   644 CCAAAAGGATAACTTTCGTATCGGCTTTGGCACTTCAGACAAGCTGAGAACCAAGGATAG   703

Qy   212 SerProValLeuLeuAspGlnAlaPheAspLeuTyrCysArgGlyValIleIleGlyPheGly   231
            |||||||:::                 :::            :::
Db   704 GGAAGCAACTCTCTTGAAGAGACCTTCGATAAGTTACCACAGGGCGTGAGTTGTATGGA   763

Qy   232 ProPheTrpGluHisMetLeuGlyTyrTrpArgGluSerLeuLysArgProGlyLysVal   251
            |||||||:||:|||||||:|||||||||||::|||||:||:           |||||
Db   764 CCGTTTTGGGACCATGTTTATGGTATTGGAAACAAGCTTGGAAAAACCTGAAAGAGTC   823

Qy   252 PhePheLeuArgTyrGluAspLeuLysAspAspPheAspIleThrAsnLeuLysArgLeuAla   271
            |||||||||||:||||||||||||||:::            |||||
Db   824 TTCTTTCTTCGAAATACGAGGACATGAAGGATGACCCACAGCCTGATGAAGCTAGCT   883

Qy   272 ThrPheLeuGluLeuProPheThrGluGluAspIleLysGlyLeuGluGluGluGlyGluAla   291
            |||||       :::            |||||
Db   884 CAGTTCCTCGGATGCCCATTTTCCAATGAAGAAGACTAGATGTGCGGTGGGCGTGTATA   943

Qy   292 AlaGluLeuCysSerSerPheGluAsnLeuLysGlyValLeuAsnLysSerAsnLysSer   311
            :::|||||||||||||||||||||||:||::            :::
Db   944 CAAAAACTATGCAGCTTTGAGAATTGACCAACTTGAGCGTTAACAATACCCGGGAAGTTG  1003

Qy   312 IleLeuAsnPheGluAsnArgPheLeuPheArgLysGlyGluValSerAspTrpValAsn   331
            :::|||||||||||||||:::            :::            :::
Db  1004 GCATCGGGTGAGGATTACAAGGGCATTTTTCACACAACGTGAAAATTGGAGAATCCTAACAAT  1063

Qy   332 Tyr   332
            :::
Db  1064 CAT  1066
```

RESULT 12

ADA68055

ID   ADA68055 standard; DNA; 980 BP.
XX
AC   ADA68055;
XX
DT   20-NOV-2003   (first entry)
XX
DE   Arabidopsis thaliana gene, SEQ ID 310.
XX
KW   Plant; bacterial infection; fungal infection; viral infection; ds.
OS   Arabidopsis thaliana.
XX
XX
PN   WO2003000898-A1.
XX
PD   03-JAN-2003.
XX
PF   22-JUN-2001; 2001WO-IB001105.
XX
PR   22-JUN-2001; 2001WO-IB001105.
XX
PA   (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI   Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI   Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR   WPI; 2003-175290/17.
XX
PT   Identifying at least one gene involved in plant resistance or response to
PT   pathogenic infection for conferring resistance or tolerance to a plant to
PT   bacterial, fungal or viral infection by determining or detecting plant
PT   gene expression.
XX
PS   Claim 6; SEQ ID NO 310; 899pp; English.
XX
CC   The present invention relates to a method (M1) for identifying genes
CC   involved in plant resistance or response to pathogenic infection. M1
CC   comprises identifying a gene whose expression is significantly altered in
CC   the incompatible interaction of plant gene expression relative to
CC   expression of the gene in an uninfected plant, in a mutant plant that

```
CC   does not express a gene associated with response to pathogenic infection,
CC   or in a corresponding incompatible or compatible interaction. (M) is
CC   useful for conferring resistance to resistance or tolerance to a plant to
CC   bacterial, fungal or viral infection. The present sequence was used to
CC   illustrate the invention.
XX
SQ   Sequence 980 BP; 282 A; 178 C; 228 G; 292 T; 0 U; 0 Other;
DB:

US-10-019-931-3 (1-359) x ADA68055 (1-980)

Alignment Scores:
Pred. No.:
Score:                    1.96e-71          Length:        980
Percent Similarity:        757.00           Matches:       148
Best Local Similarity:     61.11%           Conservative:   50
Query Match:               45.68%           Mismatches:    120
DB:                        39.88%           Indels:          6
                                            Gaps:            3

Qy   33 ValProLysAlaGluAspGluGlyGlyLeuSerCysGluPheGlnGluMetLeuAspSer 52
        |||||      ||
Db   19 GTTCCTGCTTACTTGGGAGATGAAGATCTGACACAAGAAACAAGAGCTCTGATCTCTTCT 78

Qy   53 LeuProLysGlyLeuArgGlyTyrProArgThrArgThrTyrLeuThrLeuPheGlyTrpCys 72
        |||||:|
Db   79 CTTCCTAAAGAAGGAAAAGGTTGGTTAGTGAGTGAAATATATGAAATTCCAAGGACTTGGCAC 138

Qy   73 GlnAlaLysGlyIleGlnIleAlaIleIleMetSerPheGlnHisPheGlnSerLeuGluAsn 92
        ::|||
Db  139 ACACAAGCTATTTACAAGGAATCTTGATCTGCCAAAAAAGCTTTGAAGCTAAAGAATTCC 198

Qy   93 AspValValLeuAlaThrIleProLysSerGlyThrThrThrProLeuGlyAlaLeuThrPhe 112
        |||||||
Db  199 GACATTATCCTCGTCACTAAATCCTAAAATCAGGTTACCCACTTGGTTAAAAGGTCTCTTTGTCTTT 258

Qy  113 ThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThrAsnHisProLeuPhe 132
        |||||||||||
Db  259 GCTCTCCTTAACCGACACAAGTTT---CCAGTTTCTTCTTCTGGTAACCATCCTCTTCTG 315

Qy  133 ThrSerAsnProHisAspLeuValProPhePheGluTyrTyrLeuTyrAlaAsnGlyAsp 152
        ::|||||||
Db  316 GTCACCAATCCACACCTTCTTGTGCCCTTCTTGGAAGGAGTTTACTACGAGTCCCCAGAT 375

Qy  153 ValProAspLeuSerGlyTyrLeuAlaSerProArgThrPheAlaThrHisLeuProPheGly 172
        ||||||
Db  376 TTC---GATTTCTCCAGTTTGCCTTCTCCAAGACTGATGAACACGCACATATCGCATCTT 432

Qy  173 SerLeuLysGlyThrIleGluLysProGlyValLysValValTyrLeuCysArgAsnPro 192
        |||||::||
Db  433 TCGGCTCCCGAGTCGTTGTTAAGAGACTCCGTCTTGTTGTTATTGTTGTTGTAGGAACCCT 492

Qy  193 PheAspThrPheIleSerSerThrPheHisTyrThrAsnAsnIleLysSerGluSerValSer 212
        |||:|||
Db  493 AAGGACATGTTTGTGTCCTTATGGCATTTGGGAAAAGCTAGCTCCTAGGAGAAACCCCG 552

Qy  213 ProValLeuLeuAspGlnAlaPheAspLeuTyrCysArgGlyValIleGlyPheGlyPro 232
        ||||||||
Db  553 GATTATCCTATCCGAAAAAGCGGTTGAAGGGTTTGTGAGGGAAGTTTATAGGTGGACCC 612

Qy  233 PheTrpGluHisMetLeuGlyTyrTrpArgAlaSerLeuLysArgProGlyLysValPhe 252
        |||
Db  613 TTTTGGGATCATATAATTGGAGTACTGGGATTGCAAGCCCGAGAATCCCAACAAGGTCTG 672

Qy  253 PheLeuArgTyrGluAspLeuLysAspAspIleGluThrLeuLysLysArgLeuAlaThr 272
        |||||::|
Db  673 TTTGTTACTTACGAGGACTTAAAAGATGATCAGAAGAGACCGAGAAGCGGGATCGCGAG 732

Qy  273 PheLeuGluLeuProPheThrGlnGluGluLeuArgLysGlyValValLysAlaIleAla 292
        |||||::||
Db  733 TTCTTGGAATGTGGCTTTATTGAAGAAGAAGAA------GTGAGAGATTGTG 780

Qy  293 GluLeuCysSerPheGluAsnLeuLysLysGluValAsnLysSerAsnLysSerIle 312
        :::|||  :::|||
Db  781 AAGTTGTGTAGCTTTGAGAGTTTAAGTAATTTGAAGAAGAGACGGAAATTGCCA 840
```

```
Qy  313 LysAsnPheGluAsnArgPheLeuPheArgLysGlyLeuGluValSerAspTrpValAsnTyr 332
        |||  :::          |||||:|||::|        :::          |||
Db  841 AATGGAATAGAGAGACTAAAACTTTCTTTAGAAAAGGGACTTGAAGTGAGTGATTGGGTGAATTAT 900

Qy  333 LeuSerProSerGlnValGluArgLeuSerAlaLeuValAlaAspAspLysLeuGlyGlySer 352
        |||||||::|||          |||::|||||:||||||:|        ::::|||
Db  901 TTGAGTGAGTCATTGGCAGAGGGAAAATTGATTAGAACCATTGAAGAAGAGAAGTTTAAAGGTTCT 960

Qy  353 GlyLeuThrPhe 356
        |||||||  |||
Db  961 GGTCTTAAATTT 972
```

```
RESULT 13
ID   AB214244
XX
AC   AB214244;
XX
DT   21-JAN-2003 (first entry)
XX
DE   Arabidopsis thaliana stress regulated gene SEQ ID NO 2049.
XX
KW   Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   WO020216655-A2.
XX
PD   28-FEB-2002.
XX
PF   24-AUG-2001; 2001WO-US026685.
XX
PR   24-AUG-2000; 2000US-02278866P.
PR   26-JAN-2001; 2001US-02564647P.
PR   22-JUN-2001; 2001US-0300111P.
XX
PA   (SCRI ) SCRIPPS RES INST.
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Harper JF, Kreps J, Wang X, Zhu T;
XX
DR   WPI; 2002-304127/34.
XX
PT   Identifying a stress condition to which a plant cell has been exposed and
PT   producing plants with increased tolerance to these abiotic stresses.
XX
PS   Claim 144; SEQ ID NO 2049; 577pp + Sequence Listing; English.
XX
CC   The invention relates to identifying a stress condition to which a plant
CC   cell has been exposed, comprising: (a) contacting nucleic acid
CC   representative of expressed polynucleotides in the plant cell with an
CC   array or probes representative of the plant cell genome; and (b)
CC   detecting a profile of expressed polynucleotides in the plant cell
CC   characteristic of a stress response. The method is useful in the
CC   production of transgenic plants, cells and seeds and in producing plants
CC   with increased tolerance to abiotic stress. The present sequence is that
CC   of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
CC   in methods of the invention. Note: The sequence data for this patent is
CC   not represented in the printed specification but is based on sequence
CC   information supplied to Derwent by the European Patent Office
XX
SQ   Sequence 981 BP; 283 A; 178 C; 228 G; 292 T; 0 U; 0 Other;

US-10-019-931-3 (1-359) x AB214244 (1-981)

Alignment Scores:
Pred. No.:
Score:                    1.97e-71          Length:        981
Percent Similarity:        757.00           Matches:       148
Best Local Similarity:     61.11%           Conservative:   50
Query Match:               45.68%           Mismatches:    120
DB:                        39.88%           Indels:          6
                                            Gaps:            3
```

```
Qy    33 ValProLysGlyAlaGlyGluGluAspGluGlyLeuSerCysGluPheGlnGlyMetLeuAspSer 52
            |||:::  |||   ::::::  |||     ::::::  ||||
Db    19 GTTCCTGCTTACTTGGGAGATGAAGAGAACTGACACAAGAAACAAGAGACTTGATCTCTTCT 78

Qy    53 LeuProLysGluArgGlyTrpThrArgTyrArgLeuPheGlnGlyPheThrCys 72
            |||:::||:|||  :::::|||   :::::   ||
Db    79 CTTCCTAAAGAGAGAAGGTTGGTACGGTGAGAAATATATGAATTCCAAGGACTTTGGCAC 138

Qy    73 GlnAlaLysGluIleGlnAlaIleMetSerPheGlnLysHisPheGlnSerLeuGluAsn 92
            |||    :::|||   |||    ::::::|||:::
Db   139 ACACAAGCTATTTTACCAAGGAAATCTTGATCTGCCAAAAACCGCTTTGAAGCTAAAGATTCC 198

Qy    93 AspValValLeuAlaThrIleProLysSerGlyTyrThrThrTrpLeuLeuLeuThrPhe 112
            |||:::|||   :::::|||   ||||:::
Db   199 GACATTATTCTCCGTACCACTAATCCTAAAATCAGGTACCACTGGTTAAAAGTCTTGTCTTT 258

Qy   113 ThrIleLeuLeuAsnArgHisArgPheAspProValAlaSerSerThrAsnHisProLeuPhe 132
            |||:::|||   :::::|||    :::::
Db   259 GCTCTCCTTACCGACACAAGTTT---CCAGTTTCTTCTTCTTCGGTAACCATCCTTTCTG 315

Qy   133 ThrSerAsnProHisAspLeuValProPheGluTyrGlyLeuTyrValAlaAsnGlyAsp 152
            |||:::|||   :::::  ||
Db   316 GTCACCAATCCACACCTTCTTGTGCCCTTCTTGGAAGGAGTTTACTACGAGTCCCCAGAT 375

Qy   153 ValProAspLeuSerGlyLeuValAlaSerProArgThrPheAlaThrHisLeuProPheGly 172
            |||  :::
Db   376 TTC---GATTTCTCCAGTTTGCCTTCTCCAAGACTGTGATGAACACGCACATATCGCATCTT 432

Qy   173 SerLeuLysGlyThrIleGlyLysProGlyValLysValValThrTyrLeuCysArgAsnPro 192
            |||:::|||   :::::|||
Db   433 TCGCTCCCCGGAGTCTGTTAAGAGCTGGTGTTAAGGTGGTTATTATTTGTTCGTGAGAACCCT 492

Qy   193 PheAspThrPheIleSerSerTrpHisTyrThrAsnAsnLeuLysSerGluSerValSer 212
            |||:::|||   :::
Db   493 AAGGACATGTTTGTCCTTAGCTAGCTCCTTGAGAAAACCGCG 552

Qy   213 ProValIleLeuAspGlnAlaPheAspLeuTyrCysArgGlyValIleGlyPheGlyPro 232
            |||||   :::::|||
Db   553 GATTATCCTATCGAAAAGGGTTGTGAACGGGAGTTTATNGGTGGACCC 612

Qy   233 PheTrpGluHisMetLeuGlyTyrTrpTrpArgGluSerLeuTyrLysGlyLysValPhe 252
            |||   :::|||:::  ||
Db   613 TTTTGGGATCATATTATTGGGTACTGGTGGCGGGAATCCGAACAAGGTCTTG 672

Qy   253 PheLeuArgTyrGluAspLeuThrAspGluIleGluThrArgAsnLeuLeuLysGluLeuAlaThr 272
            |||:::|||   :::::|||
Db   673 TTTGTTACTTACGAGGACCTAAAGAACGAACAGCAGAGTTGAGATGAAGCCGATCCGGGAG 732

Qy   273 PheLeuGluLeuProPheThrGluGlyLeuArgLysGlyValValValLysAlaIleAla 292
            |||:::|||      ------
Db   733 TTCTTGGAATTGGGCTTTATTGAAGGAAGAGA------GTGAGAGAGATTGTG 780

Qy   293 GluLeuCysSerProPheGluAsnLeuGlyValLeuAsnLysSerAsnLysSerIle 312
            :::|||   :::::|||
Db   781 AAGTTGTGTAGCTTTGAGAGTTAAGTAATTTGGAAGTTAACAAAGAAGGAAATGCCA 840

Qy   313 LysAsnPheGluAsnArgPheLeuPheArgLysGlyGlnValSerAspTrpValAsnTyr 332
            |||   :::|||:::  ||
Db   841 AATGGAATAGGAGACTAAAAACTTTCTTTAGAAAAGGAGGATGGAGAGATACT 900

Qy   333 LeuSerProSerGlnValArgLeuSerLeuValAspAspLysLeuGlySer 352
            |||   :::
Db   901 TTGAGTGAGTCATGGGCAGAGAAATTGATAGAACCATTGAAGAGAAGTTTCT 960

Qy   353 GlyLeuThrPhe 356
            |||||
Db   961 GGTCTTAAATTT 972
```

Alignment Scores:
Pred. No.:   1.97e-71   Length: 981
Score: 757.00   Matches: 148
Percent Similarity: 61.11%   Conservative: 50
Best Local Similarity: 45.68%   Mismatches: 120
Query Match: 39.88%   Indels: 6
DB: Gaps: 3

```
US-10-019-931-3 (1-359) x ADG87892 (1-981)

Qy      33 ValProLysAlaGlyGluGlyLeuSerCysGluPheGlnGluMetLeuAspSer 52
           |||||||
Db      19 GTTCCTGCTTACTTGGGGAGGATGAAGATCTGACAGAAGAAACAAGAGCTCTGATCTCTTCT 78

Qy      53 LeuProLysGluArgGlyTrpProArgThrArgTyrLeuPheGlnGlyPheTrpCys 72
           ||||||||||||||
Db      79 CTTCCTAAAGAGAAAAGGTTGGTTAGTGAGTGAAAATATATGAAATTCCAAGGACTTTGGCAC 138

Qy      73 GlnAlaLysGluIleGlnIleAlaIleMetSerPheGlyHisHisPheGlnSerLeuGluAsn 92
           :||||| ||||| ||||||
Db     139 ACACAAGCTATTTTACAAGGAATCTTGATCTGCCAAAAACGGCTTTGAAGCTAAAGATTCC 198

Qy      93 AspValValLeuAlaThrIleProLysSerGlyThrThrTrpLeuValAlaLeuThrPhe 112
           ||||||||||||
Db     199 GACATTATTCCTCGTCACTAATCCTAAATCAGGTACCACTTGGTTAAAAGCTCTTGTCTTT 258

Qy     113 ThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThrAsnHisProLeuPhe 132
           ::||||||||||||
Db     259 GCTCCTCCTTACCGACACAAGTTT---CCAGTTTCTTCTTCTCTCGGTAACCATCCTCTCTG 315

Qy     133 ThrSerAsnProHisAspLeuValProPheGluTyrLysGlyLeuTyrArgAlaAsnGlyAsp 152
           :::||||||
Db     316 GTCACCAAATCCACACACCTTCTTGTGCCCTTCTTGGAGGAGTTTACTACGAGTCCCCAGAT 375

Qy     153 ValProAspLeuSerGlyLeuValAlaSerProArgThrPheAlaThrHisLeuProPheGly 172
           ||||||||||
Db     376 TTC--GATTTCTCCAGTTTGCCTTCTCCCAGACTGATGAACACGACACATACGCATCTT 432

Qy     173 SerLeuLysGluThrIleGluLysProGlyValLysValValTyrLeuCysArgArgAsnPro 192
           ::||||||||||
Db     433 TCGCTCCCCCGAGTCTGTTAAGAGCTCGTCTGTTAAGGATTGTGTATTGTTGTTAGGGACCCT 492

Qy     193 PheAspThrPheIleSerSerTrpHisTyrThrAsnAsnIleLysSerGluSerValSer 212
           ||||||::||||| :::
Db     493 AAGGACATGTTTGTGTCCTTATGGCATTTTGGGAAAAAAGCTAGCTCCTTGAGAAACCGCG 552

Qy     213 ProValLeuLeuAspGlnAlaPheHisAspLeuTyrCysArgGlyValIleGlyPhePro 232
           ::||||| ::: |||
Db     553 GATTATCCTATCGAAAAAGCGGTTCGAAGCGTTTTTGTGAAGGGAAGTTTATAGGTGGACCC 612

Qy     233 PheTrpGluHisMetLeuGlyThrTyrTrpArgGluSerLeuLysArgProGlyValPhe 252
           ||||||:::|||||
Db     613 TTTTGGGATCATATATTGGAGTACTGGTATGCAAGCCCGAGAATCCGAACAAGGTCTTG 672

Qy     253 PheLeuArgTyrGluAspProLeuLysAspAlaIleGluThrArgAsnLeuLysArgLeuAlaThr 272
           :::||||||||| :::||||
Db     673 TTTGTTAGTTACGAGGAGCTAAAGAGACGACCGAGTTGAGATGAAGCCGAGATCGCGGAG 732

Qy     273 PheLeuGluLeuProPheThrGluGlyLeuGluGlyArgGlyValValLysAlaIleAla 292
           ||||| :::|||
Db     733 TTCTTGGAAATGGGCTTTATTGAAGAAGAAGAA 780

Qy     293 GluLeuGlySerPheGluAsnLeuLysLysLeuGlyValLeuAsnLysSerSerIle 312
           ::||||||||||||
Db     781 AAGTTGTTGTAGCTTTGAGAGTTTAAGTTAATTTCGGAATAGAGGGAAAATTGGCCA 840

Qy     313 LysAsnPheGluAsnArgPheLeuProArgLysGlyValSerAspTrpValAsnTyr 332
           :::|||
Db     841 AATGGAATAGAGACTAAAACTTTCTTTAGAAAAGGAGAGATTTGGAGAGGAGAGATACT 900

Qy     333 LeuSerProSerGlnValGluLeuSerAlaLeuValAlaAspAlaProLysLeuGluGlyGlySer 352
           :::||| |||
Db     901 TTGAGTGAGTCAATGGGCAGGAAAATTGATAGAACCATTGAAGAAGAAGTTTAAAGGTTCT 960

Qy     353 GlyLeuThrPhe 356
           |||
Db     961 GGTCTTAAATTT 972
```

```
XX
AC      ADG87829;
XX
DT      22-APR-2004  (first entry)
XX
DE      A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #271.
XX
KW      Pathogen infection-related gene; plant; Peronospora parasitica;
KW      defence mechanism; RPP7; RPP8; pathogen resistance; transgenic plant;
KW      oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX
OS      Arabidopsis thaliana.
XX
PN      WO20022675-A2.
XX
PD      21-MAR-2002.
XX
PF      14-SEP-2001; 2001WO-US028506.
XX
PR      15-SEP-2000; 2000US-0232778P.
PR      22-JUN-2001; 2001US-0300183P.
XX
PA      (SYGN ) SYNGENTA PARTICIPATIONS AG.
PA      (UYNC-) UNIV NORTH CAROLINA.
PA      (GLAZ/) GLAZEBROOK J.
PA      (WANG/) WANG X.
PA      (DANG/) DANGL J L.
PA      (EULG/) EULGEM T.
PA      (ZHUT/) ZHU T.
XX
PI      Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;
XX
DR      WPI; 2002-292409/33.
XX
PT      Novel isolated polynucleotide, useful for conveying pathogen resistance
PT      to plants, and for identifying plants infected with a pathogen.
XX
PS      Claim 3; SEQ ID NO 271; 605pp; English.
XX
CC      The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC      ADG87557)) whose expression is altered in response to pathogen infection,
CC      and to homologues of these genes from other plants or fungi, especially
CC      from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC      cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
CC      expression of genes of the invention was upregulated or downregulated in
CC      Arabidopsis plants infected with the oomycete Peronospora parasitica,
CC      indicating that they play a role in defence mechanisms. The genes of the
CC      invention are regulated by RPP7 or RPP8 which act via unconventional
CC      signalling cascades, or by the RPP4-dependent pathway. The invention also
CC      relates to polypeptides encoded by the pathogen infection-related genes;
CC      promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC      ; expression cassettes, host cells and pathogen-resistant transgenic
CC      plants and their progeny comprising a polynucleotide of the invention;
CC      and a method of identifying a plant cell infected with a pathogen. The
CC      polynucleotide sequences and methods of the invention are useful for
CC      identifying plants infected with a pathogen, and for conferring
CC      resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC      neamatodes and insects (e.g., aphids). The present sequence represents an
CC      Arabidopsis thaliana gene whose expression is altered in response to
CC      Peronospora parasitica infection. Note: The sequence data for this patent
CC      did not form part of the printed specification, but was obtained in
CC      electronic format directly from WIPO at
CC      ftp.wipo.int/pub/published_pct_sequences.
XX
SQ      Sequence 981 BP; 283 A; 178 C; 228 G; 292 T; 0 U; 0 Other;
```

```
RESULT 15
ADG87829
ID  ADG87829 standard; cDNA; 981 BP.
```

Alignment Scores:
```
Pred. No.:              1.
Score:                  757.00
Percent Similarity:     61.11%
Best Local Similarity:  39.88%
Query Match:            6
```
```
1.97e-71
Length:          981
Matches:         148
Conservative:    50
Mismatches:      120
Indels:          6
Gaps:            3
DB:
```

```
US-10-019-931-3 (1-359) x ADG87829 (1-981)

Qy     33 ValProLysAlaGluGluAspGluGlyLeuSerCysGluPheGlnGluMetLeuAspSer 52
          |||||||||||||||||||:::||             ||:::||||||
Db     19 GTTCCTGCTTACTTGGGAGAGATCTGACAAGAAACAAGAGAGCTCTGATCTCTTCT 78

Qy     53 LeuProLysGluArgGlyThrArgThrArgTyrLeuPheGlyPheTrpCys 72
          |||||||||||:::||||||||||||||||:::|||||
Db     79 CTTCCTAAAGAAGGAAAAGGTTGGTTAGTGAGTGAAATATATGAATTCCAAGGACTTTGGCAC 138

Qy     73 GlnAlaLysGlyIleGlnAlaIleMetSerPheGlyHisIleSerLeuGluAsn 92
          ||||||||||||||||||||:::||              ::::::
Db    139 ACACAAGGCTATTTACAGAGGAATCTTGATCTGCCAAAAACCGCTTTGAAGCTAAAGATTCC 198

Qy     93 AspValValLeuAlaThrIleProLysSerGlyThrThrTrpLeuValLysLeuAlaLeuThrPhe 112
          |||:::||||              ||:::||||||
Db    199 GACATTATCCGTCACTAATCCTAAATCAGGTACCCACTTGGTTAAAGCTTTGTCTTTTT 258

Qy    113 ThrIleLeuAsnArgHisArgPheAspProValAlaLaSerSerThrAsnHisProLeuPhe 132
          |||:::||||||:::||||||||:::||||||
Db    259 GCTCCTCCTTAACCGACACAAGTTT---CCAGTTTCTTCTTCTTCTGGTAACCATCCTTTCTG 315

Qy    133 ThrSerAsnProHisAspLeuValProPheGluTyrLysLeuTyrAlaAlaAsnGlyAsp 152
          |||||||||||:::||||||||||||||||||:::|||
Db    316 GTCACCAAATCCACACCTTCTTGTGCCCTTCTTGGAGGAGTTCCCCAGAT 375

Qy    153 ValProAspLeuSerGlyLeuAlaSerProArgThrPheAlaThrHisLeuProHisGly 172
          |||||              ||:::||||||
Db    376 TTC---GATTTCTCCAGTTTGCCTTCTCCAAGACTGATGACACGCACATAACGCCATCTT 432

Qy    173 SerLeuLysGluThrIleGluLysProGlyValLysValValTyrLeuCysArgAsnPro 192
          |||:::||||||:::|||||||
Db    433 TCGCTCCCCGAGTCTGTTAAGAGCTCGTCTGTTCGTCTTGTAATGTGTTGTTGTTGGAACCT 492

Qy    193 PheAspThrPheIleSerSerTrpHisTyrThrAsnAsnIleLysSerGluSerValSer 212
          |||              ||:::
Db    493 AAGGACAGTTTGTGTCCTTATGGCATTTTGGGAAAAAGCTAGCTCCTGAGAAACCGCG 552

Qy    213 ProValLeuLeuAspGlnAlaPheAspGluLeuTyrCysAgrGlyValIleGlyPheGly 232
          |||||              ::: ||
Db    553 GATTATCCTATCGAAAAAGCCGTTGAAGCGTTTTTGCAAGGGAGTTTATAGGTGGACCC 612

Qy    233 PheTrpGluHisMetLeuGlyTyrrTpgArgGluSerLeuLysSerArgProGluLysValPhe 252
          |||||||:::|||:::||||              ::: ||
Db    613 TTTTGGGATCATATATTGGAGTACTGGTATGCCAAGCCCGACAAAGGTCTTG 672

Qy    253 PheLeuArgTyrGluAspLeuLysSerAspIleGluThrArgAsnLeuLysSerArgLeuAlaThr 272
          |||:::              |||:::|||
Db    673 TTTGTTACTTACGAGGAGCTAAAGAGAAGCAGAACCGGATTGAGATGAAGCCGGAG 732

Qy    273 PheLeuGluLeuProPheThrGluGluGluGluLysGlyValValLysValLysAlaIleAla 292
          ||||||:::              |::: |||
Db    733 TTCTTGGAATGTGGCTTTTATTGCAAGGAAGAAA---GTGAGAGAGATTGTG 780

Qy    293 GluLeuCysSerPheGluAsnLeuGlyLeuGluValValAsnLysSerAsnLysSerSerIle 312
          ::::::|||              |||
Db    781 AAGTTGTGTAGCTTTGAGAGTTTAAGTGTAATTTGGAAGTTAACAAAGAGGGAAAATGCCA 840

Qy    313 LysAsnPheGluAsnArgPheLeuArgLysGlyGluValSerSerAspTrpValAsnTyr 332
          |||:::|||              |||:::
Db    841 AATGGAATAGAGACTAAAAACTTTCTTTAGAAGAATTTGGAGAGTGGAGAGATACT 900

Qy    333 LeuSerProSerGlnValGluArgLeuSerAlaLeuValAlaAspProLysLeuGlyGlySer 352
          ||||||:::              ::::::
Db    901 TTGAGTGAGTCATTTGGCAGAGGAAAATTGATAGAAACCATTGAAGAGAAGTTTAAAGGTTCT 960

Qy    353 GlyLeuThrPhe 356
          |||:::
Db    961 GGTTCTTAAATTT 972
```

OM protein - nucleic search, using frame_plus_p2n model

Copyright (c) 1993 - 2005 Compugen Ltd.

GenCore version 5.1.6

Run on: December 23, 2005, 18:49:27 ; Search time 3717 Seconds
(without alignments)
4518.849 Million cell updates/sec

Title:　　　　US-10-019-931-3
Perfect score: 1898
Sequence: 1 MATSSMKSIPMAIFSFSMCH.........RLSALVDDKLGGSGLTPRLS 359

Scoring table: BLOSUM62
　　Xgapop 10.0 , Xgapext　0.5
　　Ygapop 10.0 , Ygapext　0.5
　　Fgapop　6.0 , Fgapext　7.0
　　Delop　　6.0 , Delext　　7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
　　　　　　　　　Maximum Match 100%
　　　　　　　　　Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model/ -DEV=xlp
-Q=/cgn2_1/USPTO_spool_P/US10019931/runat_16122005_165219_18452/app_query.fasta_1.519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10019931 @CGN 1 1 8010 @runat 16122005 165219 18452 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	1877	98.9	1396	4	CNS0A0WE	BX832510	Arabidops
2	1791.5	94.4	1201	4	CNS0A2ZK	BX831973	Arabidops
3	1778	93.7	1456	4	CNS0A0VN	BX831965	Arabidops
4	1582	83.4	1517	4	CNS0A186	BX829897	Arabidops
5	1573	82.9	1320	4	CNS0A1K2	BX829446	Arabidops
6	1533	80.8	1302	4	CNS09YST	BX831024	Arabidops
7	1528	80.5	1328	4	CNS0A0WP	BX831880	Arabidops

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

	No	Score	%		Len		Accession	Accession / Clone
c	8	1103.5	58.1	9	757	9	BZ062934	BZ062934 11c02b11.
	9	1094	57.6	8	811	9	BH733391	BH733391 BOMIU60TR
	10	968	51.0	6	727	6	CD837475	CD837475 BN45.052G
	11	936.5	49.3	5	734	6	CD834047	CD834047 BN45.040G
	12	916.5	48.3	1	671	1	AU238005	AU238005 X9SP07808
	13	902	47.5	7	875	7	CV130872	CV130872 UCRCS07_1
	14	887.5	46.8	7	817	8	CX043920	CX043920 UCRCS07_1
	15	870	45.8	6	574	1	AV827990	AV827990 AV442770
	16	854	45.0	1	523	1	CD833738	CD833738 BN45.001G
	17	853	44.9	8	651	6	CX046401	CX046401 UCRCS07_7
	18	847.5	44.7	8	717	8	CN189921	CN189921 UCRCS06_0
	19	830.5	43.8	7	782	7	CB260303	CB260303 81-E9537-
	20	819	43.2	6	657	6	AV442770	AV442770 AV442770
	21	804	42.4	1	641	1	AV827990	AV827990 AV827990
c	22	798	42.0	1	638	1	CX043920	CX043920 UCRCS07_1
c	23	798	42.0	1	585	6	AV520946	AV520946 AV520946
c	24	797.5	42.0	7	808	7	CD822905	CD822905 BN25.047A
c	25	797	42.0	9	760	9	CB260353	CB260353 77-E01117
c	26	797	42.0	6	795	6	AV441110	AV441110 AV441110
c	27	790.5	41.6	1	795	6	CN782243	CN782243 EST00339
c	28	785.5	41.4	1	531	1	CC952179	CC952179 BOIGS48TF
c	29	777	40.9	1	738	6	CA795562	CA795562 sat35e10.
c	30	770.5	40.6	7	781	6	AV442439	AV442439 AV442439
c	31	757	39.9	8	889	7	CF418052	CF418052 sat54b11.
c	32	755	39.8	4	1274	4	CA782758	CA782758 USDA-FP_1
c	33	750.5	39.5	5	918	5	CV236747	CV236747 WS01225.B
c	34	745	39.3	5	698	5	CNS0A9V7	Arabidops
c	35	733.5	38.6	8	847	7	DR927927	Arabidops
c	36	729.5	38.4	8	849	7	BX819152	Arabidops
c	37	727.5	38.3	10	934	10	BU831310	BU831310 T020A09 P
c	38	727	38.2	5	701	6	DR927927	DR927927 EST111946
	39	724	38.1	5	694	5	CO072215	CO072215 GR__Ea380
	40	719	37.9	8	855	10	CO072215	CO072215 GR__ED024
	41	716	37.7	6	680	6	CO120965	CO120965 GR__ED024
	42	707.5	37.3	7	719	8	CG919663	CG919663 MBEBT65TR
	43	681	35.9	8	916	8	BU834467	BU834467 T061E09_P
	44	680.5	35.9	10	999	10	CF503356	CF503356 USDA-FP_1
	45	678.5	35.7	8	832	8	CG950535	CG950535 MBEGOO3TF

(continued) CB006702 VVC03H04 ; DT043830 Mdrtp1007 ; CX547767 gmtdrNS0 ; CL954941 OsIRVA005 ; CX192873 38-E02278

RESULT 1
CNS0A0WE

LOCUS   CNS0A0WE   1396 bp   mRNA   linear   HTC 06-FEB-2004
DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone
            GSLTPGH78ZA04 of Hormone Treated Callus of strain col-0 of
            Arabidopsis thaliana (thale cress).
ACCESSION   BX832510
VERSION   BX832510.1 GI:42458247
KEYWORDS   HTC; GSLT_cDNA.
SOURCE   Arabidopsis thaliana (thale cress)
  ORGANISM Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1 (bases 1 to 1396)
  AUTHORS   Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
            Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
            Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
  TITLE   Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
            A Combined Approach to Evaluate and Improve Arabidopsis Genome
            Annotation
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 1396)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr :
            - Web : www.genoscope.cns.fr)
COMMENT   The sequences are based on single pass reads.
            Life Technologies (a division of Invitrogen) members carried out
            full-length librairies construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

```
FEATURES             Location/Qualifiers
source               1..1396
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="GSLTPGH78ZA04"
                     /tissue_type="Hormone Treated Callus"
                     /ecotype="Col-0"
                     /plasmid="pCMVSPORT 6"
gene                 1..1396
                     complement(1..1396)
                     /gene="At5g07010"
ORIGIN
```

```
Alignment Scores:
Pred. No.:               1.4e-204
Score:                   1877.00
Percent Similarity:      100.00%
Best Local Similarity:   99.72%
Query Match:             98.89%
DB:                      4

Length:       1396
Matches:      354
Conservative: 1
Mismatches:   0
Indels:       0
Gaps:         0
```

US-10-019-931-3 (1-359) x CNS0A0WE (1-1396)

```
Qy   5 SerMetLysSerIleProMetAlaIleProSerPheSerMetCysHisGlySerLeuGluLeu 24
         :::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  49 GCCATGAAGAGCATTCCAATGGCGATCCCAAGTTTCTCCATGTGTCACAAGCTCGAGCTC 108

Qy  25 LeuLysGluGlyLysThrArgAspValProLysAlaGluGluGlyTyrLeuSerCys 44
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 109 CTTAAAGAAGGCAAAAACTCCCGACCGTCCCCAAAAGCCGAAGAAGGCTAAGCTGC 168

Qy  45 GluPheGlnGlyMetLeuAspSerLeuProLysGluArgGlyThrTyrThrArgTyrLeu 64
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 169 GAGTTCCAAGGAATGTTGGATTCTCTTCCTAAGGAGAGGGGAACTTACACCCGTTACCTT 228

Qy  65 TyrLeuPheGlnGlyPheTrpCysGlnAlaLysGlyIleGlnIleAlaIleMetSerPheGln 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 229 TACCTATTCCAAGGGTTTTGGTGCCAAGGCAAAAGGAGATTCAAGCCATCAATCTTTCCAA 288

Qy  85 LysHisPheGlnSerLeuGlnAsnAspValValLeuAlaThrIleProLysSerGlyThr 104
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 289 AAACATTTCCAATCCCTCGAAAACGACGTGGTTCTCGCCACCATACCTAAATCCGGTACA 348

Qy 105 ThrTrpLeuLysAlaLeuThrProMetThrIleLeuLysAsnArgHisArgPheAspProValAla 124
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 349 ACCTGGCTAAAAGCTTTAACTTCACCATCCTTAAGAATCGGCATCGATTTTGAT 408

Qy 125 SerSerThrAsnHisProLeuPheThrSerAsnProHisAspLeuValProPheGlu 144
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 409 TCGAGTACCAACCACCCTCTTTTCACTTCCAACCCTCATGACCTTGTACCTTTCTTCGAG 468

Qy 145 TyrLysLeuTyrArgLaAsnGlyAspValAspProLeuSerGlyLeuAlaSerProArgThr 164
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 469 TACAAGCTTTACGCCAACGGAGATGTTGACCCGGTCTTAGGCCTCCAGAAACG 528

Qy 165 PheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluGlyLysProGlyValLys 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 529 TTCGCAACCCACTTACCGTTCGGTTCCCTAAAGGAAACGATCGAGGGAAAACCCGGTGTGAAG 588

Qy 185 ValValTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrpHisTyrThrAsn 204
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 589 GTCGTGTACTTGTGCCGGAACCCGTTTGACACATTCATCTCTTCGTGGCATTACACCAAC 648
```

RESULT 2

```
LOCUS       CNS0A0ZK    1201 bp    mRNA    linear    HTC   06-FEB-2004
DEFINITION  Arabidopsis thaliana Full-length cDNA Complete sequence from clone
            GSLTPGH432B05 of Hormone Treated Callus of strain col-0 of
            Arabidopsis thaliana (thale cress).
ACCESSION   BX831973
VERSION     BX831973.1    GI:42458076
KEYWORDS    HTC; GSLT_cDNA.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1 (bases 1 to 1201)
AUTHORS     Castelli,V., Aury,J.M., Wincker,P., Clepet,C., Schachter,V.,
            Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
            Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE       Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
            A Combined Approach to Evaluate and Improve Arabidopsis Genome
            Annotation
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 1201)
AUTHORS     Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     The sequences are based on single pass reads.
            Life Technologies (a division of Invitrogen) members carried out
            full-length libraries construction : Temple G.
            Genoscope members carried out sequencing and annotation : Castelli
            V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
            Schachter V., Weissenbach J., Salanoubat M.
            URGV INRA : Clepet C., Caboche M.
            Annotation is based on the June 2003 version of the Arabidopsis
            genome released by MIPS (Munich Information center for Protein
            Sequences). 5 prime and 3 prime are assembled with Phrap.
            http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
```

```
Qy 205 AsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaPheAspLeuThrTyrCys 224
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 649 AACATCAAATCCGAGTCAGTACAAGTTTGCTTGCTTGGATCAAGCCTTTTGATCTGTATTGC 708

Qy 225 ArgGlyValIleGlyPheGlyProPheTrpGluHisMetLeuGlyTyrThrpAspGluSer 244
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 709 CGGGGAGTGATCGGGTTTGGGCCGTTTTGGGAACACATGTTGGGATACTGAGAGAGAGC 768

Qy 245 LeuLysArgProGlyLysValPheProGluLysLysAspLeuLeuLysLysAspAlaIleGlu 264
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 769 TTGAAGAGAGACCCAGGAAGAAGTTCTTTTTAAGGTAAGGACGGAGGACATCGAG 828

Qy 265 ThrArgAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGluGluGluArg 284
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 829 ACCAACTTGAAGAGGCTTGCCAACTTTCCTTAGAGCTTCCTCACGAGTGGAAGAGGAGAACGA 888

Qy 285 LysGlyValValLysAlaIleAlaGluLeuCysSerSerPheGluAsnLeuLysGlyLeuGlu 304
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 889 AAGGGAGTTGTGAAGGCTATCGCCGAGCTGTGTAGCTTCGAAAATCTGAAAGGTTGGAG 948

Qy 305 ValAsnLysSerAsnLysSerIleLysAsnPheGluAsnPheLeuPheAsnArgLysGlyGly 324
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 949 GTGAACAAGTCAAACAAGTCCATCAAGAACTTTGAGAATCGATTCTTGTTCCGGAAAGGA 1008

Qy 325 GluValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArgLeuSerAlaLeu 344
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1009 GAAGTGAGTGATTGGGTTAACTATTTGTCACCTTCACAAGTGGAAAGAGTTCAGCCTTA 1068

Qy 345 ValAlaAspAspLeuGlyGlySerGlyLeuThrPheArgLeuSer 359
         |||||||||||||||||||||||||||||||||||||||||||
Db 1069 GTGGGATGACAAGTTAGGTGGATCTGGTCTAACTTTCAGGTTGAGC 1113
```

```
FEATURES          Location/Qualifiers
source            1..1201
                  /organism="Arabidopsis thaliana"
                  /mol_type="mRNA"
                  /db_xref="taxon:3702"
                  /clone="GSLTPGH413B05"
                  /tissue_type="Hormone Treated Callus"
                  /ecotype="Col-0"
                  /plasmid="pCMVSPORT 6"
gene              complement(1..1201)
                  /gene="At5g07010"

ORIGIN

Alignment Scores:
Pred. No.:
Score:              Length:         1201
                    Matches:        348
Percent Similarity: Conservative:   1
Percent Similarity: Mismatches:     10
Best Local Similarity: Indels:      3
Query Match:        Gaps:           1

US-10-019-931-3 (1-359) x CNS0A0ZK (1-1201)
```

```
Qy    1 MetAlaThrSerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHi 20
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGGCTACCTCACTAAGCATGAAGAGCATTCCAATGGCAATCCCAAGTTTCTCCATGTGTCA 93

Qy   20 slysLeuGlyLeuLeuLysGluGlyGlyLysThrArgAspVal-ProLysAlaGluGluAspG 40
        |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   94 CAAGCTCGAGGCGCCTTAAAGGAGGGAAAACTCGCCACAGTGTCCCGAAAGCCGAAGAAGATG 153

Qy   40 luGlyLeuSerCysGluPheGlnGluMetLeuAspSerLeuProLysGluLuArgGlyTrpA 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  154 AAGGGACTAAGCTCCGAGTTCCAAGAGATGTTGGATTCTCTTCCTAAGGAGAGAGGATGGA 213

Qy   60 rgThrArgThrTyrLeuTyrLeuPheGlnGlyPheThrProCys---GlnAlaLysGluGlnA 79
        |||||||||||||||||||||||||||||||||||||||||   |:|||
Db  214 GAACTCGGTACCTTACCTTATCCAAGGGTTTGGTGGGGGGGGGGGGGTGGG 273

Qy   79 laIleMetSerPheGlnLysHisPheGlnSerLeuGluAsnAspValValLeuAlaThrI 99
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  274 CCATCATGTCTTTCCAAAAACATTTCCAATCCCTCGAAAATCGTTCGTCGGCCACCA 333

Qy   99 leProLysSerGlyThrThrIleLeuLysAlaLeuThrThrIleLeuAsnArgHisA 119
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  334 TACCTAAATCCGGTACAACCTGGCTAAAAGCTTTAAACTTTCACCATACTTAACCGTCACC 393

Qy  119 rgPheAspProValAlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAspL 139
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  394 GGTTTGATCCGGTTGCCTCGAGTACCAACCACCCTCTTTTCACTTCAACCCTCATGACC 453

Qy  139 heLeuProPhePheGluTyrLysTyrLysValProAspLeuSerGlyL 159
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  454 TTGTACCTTTTCTTCGAGTACAAGCTTTACCGCGAACCCGGTTTGACACATTCATCTCTT 513

Qy  159 euAlaSerProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleG 179
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  514 TAGGCCAGTCCAAGAACGTTCGCACACCGGTTCCGGTTCCCTAAGGAAACGATACG 573

Qy  179 luLysProGlyValLysValTyrLeuCysArgAspAsnProPheAspThrPheIleSers 199
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  574 AGAAACCCGGTGAAGGTGTATCTGTGCCGGGAACCCGGTTGACACATTCATCTCTT 633

Qy  199 erTrpHisTyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnA 219
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  634 CGTGGCATTACACCAACAACATCAAAATCCCAGTCAGTTACCCAGTCTTGCTAGACCAAG 693

Qy  219 laPheAspLeuTyrCysArgGlyValAlaIleGlyLysProPheGlyProPheGluHisMetLeuG 239
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  694 CTTTTGATCTGTATTGCCGGGGAGTCGCCGGTTGGGAAGCCGTTTGGGAACACACATGGAT 753

Qy  239 lyTyrTrpArgGluSerLeuLysArgProGluValPhePheLeuArgTyrGluSerPr 259
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  753 GGATACTGGAGAGAGAGCTTGAAGAGACCAGAGAAAGTCTCTTCTTTTTAAGTTACGAGGATC 813
```

```
RESULT 3
CNS0A0VN
LOCUS       CNS0A0VN            1456 bp   mRNA   linear   HTC   06-FEB-2004
DEFINITION  Arabidopsis thaliana Full-length cDNA Complete sequence from clone
            GSLTPGH422C07 of Hormone Treated Callus of strain col-0 of
            Arabidopsis thaliana (thale cress).
ACCESSION   BX831965
VERSION     BX831965.1   GI:42458074
KEYWORDS    HTC; GSLT cDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1456)
  AUTHORS   Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
            Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
            Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
  TITLE     Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
            A Combined Approach to Evaluate and Improve Arabidopsis Genome
            Annotation
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1456)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     The sequences are based on single pass reads.
            Life Technologies (a division of Invitrogen)
            full-length libraries construction : Temple G.
            Genoscope members carried out sequencing and annotation : Castelli
            V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
            Schachter V., Weissenbach J., Salanoubat M.
            URGV INRA : Clepet C. Caboche M.
            Annotation is based on the June 2003 version of the Arabidopsis
            genome released by MIPS (Munich Information center for Protein
            Sequences). 5 prime and 3 prime are assembled with Phrap.
            http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
            length
            http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=arabidopsis.
FEATURES          Location/Qualifiers
source            1..1456
                  /organism="Arabidopsis thaliana"
                  /mol_type="mRNA"
                  /db_xref="taxon:3702"
                  /clone="GSLTPGH422C07"
```

```
Db  754 GATACTGGAGAGAGAGCTTGAAGAGACCAGAGAAAGTCTCTTCTTTTTAAGTTACGAGGATC 813

Qy  259 euLysAspAspAspIleGluThrAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheT 279
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  814 TCAAAGACGACGACATCGAGACCAACCTGAAGAGACGCCTTGCAACTTTCTTAGAGCTTCCTTCA 873

Qy  279 hrGluGluGluLysArgLysGlyGlyValValIleValAlaIleAlaGluLeuCysSerPheLuA 299
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  874 CCGGAGAAGAGAAGAACCGAAAGGGAGGGAGTGTGAAGGCTTGTGTGAGCCTTCGGAGA 933

Qy  299 snLeuLysGlyLeuGlyGluValAlaAsnLysLysSerArgSerIleIleLysLysAsnPheGluA 319
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  934 ATCTGAAGGGATTGGGTGAAGTGGCGAACAAGTCCGATCAAGAACTTTGAGAATCGAT 993

Qy  319 heLeuPheArgLysGlyGluValSerArgAspTrpValAsnTyrLeuSerProSerGlnValG 339
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  994 TCTTCGTTTCCGGAAAGGAGGAAGTGAGTGGGTAAACATATTTTGTCACCCTTCACAAGTGG 1053

Qy  339 luArgLeuSerAlaLeuValAlaAspAspLysLysLysGlySerGlyLeuThrPheArgLeuS 359
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1054 AAAGATTGTCAGCCCTTAGTGGATGACAAGTTAGGTGGATGGTGGTCTGGACTTTCAGGTTGA 1113

Qy  359 er 359
        ||
Db  1114 GC 1115
```

```
ORIGIN
gene      /tissue_type="Hormone Treated Callus"
          /ecotype="Col-0"
          /plasmid="pCMVSPORT 6"
          complement(1..1456)
          /gene="At5g07010"

Alignment Scores:
Pred. No.:
Score:                3.73e-193      Length:          1456
                      1778.00        Matches:         337
Percent Similarity:   96.90%         Conservative:    7
Percent Identity:     94.93%         Mismatches:      11
Query Similarity:     93.68%         Indels:          0
Best Local Similarity:               Gaps:            0
DB:                   4

US-10-019-931-3 (1-359) x CNS0A0VN (1-1456)
```

```
Qy   5  SerMetLysSerIleProMetAlaIleProSerPheSerMetCysHisLysLeuGluLeu  24
        ::::||:|||
Db  73  GCGGTGAAGAGAGCATTCCAATCGCGATCCCAAGTTTCTCCATGTGTCACAAGGCTC   132

Qy  25  LeuLysGluGlyLysThrArgArgSerProLysAla...                       44
        |||||||||||||||||||||||
Db 133  CTTAAAGAAGGCAAAAACTCGCCACGTTCCCAGAAAGCCCAAGAAGATGAAGGGCTAAGCTGC 192

Qy  45  GluPheGlnGluMetLeuAspSerLeuProLysGlyArgGlyTrpProArgThrArgTyrLeu 64
        |||||||||||||||||||||||||||
Db 193  GAGTTCCAAGAGATGTTGGATTCTCTTCCTAAGGGACGAGGATGGAGAACTCGTTACTT   252

Qy  65  TyrLeuPheGlnGlyPheTrpCysSerGlnIleAlaLysGluIleGlnIleAlaIleMetSerPheGln 84
        |||||||||||||||||||||||||
Db 253  TACCTATTCCAAGGGTTTTGGTGCTCAAAGATTGCAAAGAGAATTCAAGCCATCATGTCTTTCCAA 312

Qy  85  LysHisPheGlnSerLeuGluGluAsnAspValValLeuAlaThrIleProLysSerGlyThr 104
        ||||||||:::||||
Db 313  AAACATTTTCAATCCCTCGATAAGCGACGTCGTTCTCCGCCACCATACCTAAATCCGGTACA  372

Qy 105  ThrThrProLeuLysAlaLeuThrPhePheThrIleLeuAsnArgHisArgPheAspProValAla 124
        |||||||||||||||||
Db 373  ACCTCGCCTAATAGCTTTAACTTTCACCATCCTTAACCGTCACCGGTTTGATCCGGTTGCC   432

Qy 125  SerSerThrAsnHisProLeuPheThrSerArgAsnProHisAspLeuValProPhePheGlu 144
        ||||||||||||||
Db 433  TCGAGTACCAACCACCCTCTTTTCACTTCCAACCCTCATGACTTGTACTTTCTTCGAG       492

Qy 145  TyrLysLeuTyrAlaAsnGlyAspValProAspLeuSerGlyLeuAlaSerProArgThr   164
        ||||||||||||||||
Db 493  TACAAGCTTTACGCCAACGGAGATGTTCCCGATCCTCTCGGGTTAGCCTCCAGGAACG     552

Qy 165  PheAlaThrHisLeuProPheGlySerLeuLysGlyThrIleGluTyrProGlyValLys   184
        ||||||||||||||:::
Db 553  TTCGCAACCCACTTACCGTTCGGTTCCCTAAAGGAAACGATCGATAAACCCGGTGTTGAAG   612

Qy 185  ValValTyrLeuCysArgArgAsnProHeAspThrPheIleSerSerTrpHisTyrThrAsn 204
        ||||||||||||||
Db 613  GTCGTGTACTTGTGCCGGAGACCGGTTTGACACATTCATCTCTTCGTGGCATTACACCAAC   672

Qy 205  AsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaPheAspLeuTyrCys   224
        |||||||||||||
Db 673  AACATCAAAATCCGAGTTAGCCCAGTCCTTGCTAGACCAAGCTTTTTTATCTGTATTGC     732

Qy 225  ArgGlyValIleGlyProPheThrTrpGluHisMetLeuGlyTyrTrpArgGluSer       244
        ||||||||||||
Db 733  CGGGGAGTGATCGGGGTTTGGGCCCGGTTTTGGGAACACATGTTGGGATACTGGAGAGAGC   792

Qy 245  LeuLysArgProGluLysValPhePheLeuArgTyrGluAspLeuLysGlyAspAspIleGlu 264
        ||||||||||||
Db 793  TTGAAGAGACCAGAGAAGTTGTCTTCTTTTTTTAGGTACGGAGATCTCAAAGGACGACATCGAG 852

Qy 265  ThrAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGluLeuGluGluArg     284
        |||||||||||:::||
Db 853  ACCAACTTGAAGAGGCTTGCTACTTTCCTTTCACCGAAGAGATGAGGAAGAGAACGT       912
```

```
Qy 285  LysGlyValValLysAlaLeuAlaGluLeuCysSerPheGluAsnLeuLysGlyLeuGlu    304
        |||||||||||:::|||||
Db 913  GGGGGTGGTTGTGGAAGGCTATCCCGAGCTGTGTGACGTGTAGCTTCGAGAATCTGGAG      972

Qy 305  ValAsnLysSerAsnLysSerIleLysLysAsnPheGluAsnArgPheLeuPheArgLysGly 324
        ||||||||||||||||
Db 973  GTGTACAAGTCAAACAAGAGTCGATCTAGAACTTAGAACTTGAGAATTCTTGTTCGGAATGGA 1032

Qy 325  GluValSerAspTrpValAsnTyrLeuSerProSerGlnIleValGluArgLeuArgLeuSer  344
        |||||||||||||
Db 1033 GAAGTGAGTAGTTTGGGTTAACTATTTGTCACCTTCACAAGTGGAAAGAATTGTTCAGCCTTT 1092

Qy 345  ValAspAspLysLeuGlyGlySerGlyLeuThrPheArgLeuSer                   359
        |||:::|||||||||
Db 1093 GTGGATGACGAGGTTAGGTGGGATTCGGTTCTCACTTTCTGGTTGGAGC              1137
```

RESULT 4

```
CNS0A186
LOCUS       CNS0A186               1517 bp    mRNA    linear   HTC 06-FEB-2004
DEFINITION  Arabidopsis thaliana Full-length cDNA Complete sequence from clone
            GSLTFB42G02 of Flowers and buds of strain col-0 of Arabidopsis
            thaliana (thale cress).
ACCESSION   BX829897
VERSION     BX829897.1  GI:42458543
KEYWORDS    HTC; GSLT cDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1517)
  AUTHORS   Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Cleper,C.,
            Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
            Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
  TITLE     Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
            A Combined Approach to Evaluate and Improve Arabidopsis Genome
            Annotation
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1517)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     The sequences are based on single pass reads.
            Life Technologies (a division of Invitrogen) members carried out
            full-length libraries construction : Temple G.
            Genoscope members carried out sequencing and annotation : Castelli
            V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
            Schachter V., Weissenbach J. Salanoubat M.
            URGV INRA : Clepet C., Caboche M.
            Annotation is based on the June 2003 version of the Arabidopsis
            genome released by MIPS (Munich Information center for Protein
            Sequences). 5 prime and 3 prime are assembled with Phrap.
            http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
            length
            http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis
FEATURES             Location/Qualifiers
     source          1..1517
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="GSLTFB42G02"
                     /tissue_type="Flowers and buds"
                     /ecotype="Col-0"
     gene            complement(1..1517)
                     /plasmid="pCMVSPORT 6"
                     /gene="At5g07000"
ORIGIN

Alignment Scores:
Pred. No.:
Score:               1.46e-170      Length:   1517
                     1582.00        Matches:  305
```

```
Percent Similarity:     92.46%          Conservative:  26
Best Local Similarity:  85.20%          Mismatches:    21
Query Match:            83.35%          Indels:         6
DB:                                     Gaps:           3

US-10-019-931-3 (1-359) x CNS0A186 (1-1517)

Qy    4 SerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHisLysLeuGlu 23
        |||||||||::|||||||||||||||||||||||||||||||||||:|||||
Db  225 TCAAGCATAAAGAGACATTCCAATGGCGATCCCAAGTTCTCCATGTGTCACAAGCCGAG 284

Qy   24 LeuGluLysGlyGlyLysThrArgAspValProLysAlaGluGluAspGlyGlyLeuSer 43
        |||||||||||||||||:::                    ::||||||||||||||||
Db  285 CTCCTTAAGGAAGGCAAAAGC---------------GAAGGCCAAGAAGAAGAGGGCTAAGC 332

Qy   44 CysGluPheGlnGluMetLeuAspSerLeuProLysGlyLeuArgGlyGlyTrpArgTyr 63
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  333 TACGAGTTCCAAGAGATGTTGGACTCTCTTCCTAAGGAGAGAGAGGACGGAGAAATCGTTAC 392

Qy   64 LeuTyrLeuPheGlnGlyTyrPheTrpCysGlnAlaLysGlyIleIleMetSerPhe 83
        ||||||||||||||||||||||||||||||||||||||||||||||
Db  393 CTTTACTTATTCCAAGGGTTTCCGGTGCCAAGCTAAGGAGGAATTCACAGTCTTTC 452

Qy   84 GlnLysHisPheGlnSerLeuGluAsnAspValValLeuValAlaThrIleProLysSerGly 103
        ::||||||||||||||||||||||||||||||||||||||||||
Db  453 CAAAAACATTTTCAGTCCCTTCCAGACGACGTTGTCCTCGCCACCATACCTAAATCTGGC 512

Qy  104 ThrThrThrTrpLeuLysValAlaLeuThrPheThrIleLeuAsnArgHisArgPheAspProVal 123
        |||||||||||||||||||||||||||||||||||||||||||||||
Db  513 ACAACCTCGTTAAAAGCTTTAACTTTCACCATACTTACCCGTCATCGGTTTGATCCGGTT 572

Qy  124 --AlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAspLeuValProPhe 142
        ::|||||||||||||||||||||||||||||||||||||||||||||
Db  573 TCCTTCATCAAGTTCCGACCACCCTTCTTCTCACATCCAACCCTACCCGACCCTCGTACCTTTC 632

Qy  143 PheGluTyrLysLeuTyrArgAlaAsnGlyAspValProAspLeuSerGlyLeuAlaSerPro 162
        :::|||||||||||||||||||||||||||||||||||||||||||||
Db  633 TTCGAGTACAAGCTTTACGGCGCCAACGGAAATGTCCCGGATCTCTCGGGTCTTAGCCAGTCCA 692

Qy  163 ArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluLysProGly 182
        |||||||||||||||||||||||||||||||||||||||||
Db  693 AGAACATTCGCACCACCGTTCCGTCGCCTTAAGGATTCCGGTCCAGAATCCAGT 752

Qy  183 ValLysValTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrpHisTyr 202
        |||||||||||||||||||||||||||||||||||:::
Db  753 GTGAAGGTTTGTGTACCTGTGCCCGGAACCCCGTTTGACACATTCATCTCCATGTGGCATTAC 812

Qy  203 ThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaPheHisLeu 222
        |||||||||||||:::||||||||||||||||||::::|||
Db  813 ATCAACAACATCACTTCCGAGTCAGTGACGCGCAGTCTTGCTCTAGACGAAGCTTTTGATCCTA 872

Qy  223 TyrCysArgGly---ValIleGlyPheGlyProPheTrpLeuHisSerMetLeuGlyTyrTrp 241
        ||||||||||||    |||||||||||||||||||||||||||||||
Db  873 TATTGCCGCGGGATTACTGATCGGGATTTGGGAACACATGTTGGGATACTGG 932

Qy  242 ArgGluSerLeuLysArgProGluLysProGlyLysValPhePheLeuArgTyrGluAsp 261
        |||||||||||:::||||||||||||||||||:::
Db  933 AGAGAGAGCTTGAAGAGGCCCGAGAAAGTCTATTTTTAAAGTACGAGGATTCAAAGAA 992

Qy  262 AspIleGluThrAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGluGlu 281
        |||||||||||||||||||||||||||||||||||||
Db  993 GACATCGAGACCAACTTGAAGAACCTAGAAGTTTCTTAGGACGTTCCAAGAA 1052

Qy  282 GluGluLysGlyLysGlyLysValLysAlaGluLeuCysSerPheGluAsnLeuLys 301
        |||||||||||||||||||||||||||||||||||::|
Db 1053 GAGGAACAAAAGGGAGTTGTGAAAGCTAGCTACTCGCCTCGATCTGTGTAGCTTTGAGAACTGAG 1112

Qy  302 LysLeuGluLysValAsnLysSerAsnPheGluAsnArgPheLeuPhe 321
        ||||||||||||:::|||||||||||||||||||::|||
Db 1113 AAGTTGGAGAGGTGAACAAGTCAAGAAAATTGATCCAGAACTATAGAGAACCCGGTTCTTGTTT 1172

Qy  322 ArgLysGlyValValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArgLeu 341
        |||||||||||||||||||||||||||||||||||||||||
Db 1173 TCAAGCATAAAGAGACATTCCAATGGCGGATCCCAAGTTCCAGTTCCAG 341
```

Db 1173 AGGAAAAGGAGAAAGTGAGTGAGTGAGTTTGGTTAACTATTTGGTTCGCCATCACAAGTGGAAAGATTG 1232
Qy 342 SerAlaLeuValAspAspLysGlyLeuGlyGlySerGlyLeuThrPheArgLeuSer 359
Db 1233 TCAGCCTTAGTGATGACAAGTTAGTGGATTCTGGTTCTCACTTTCAGATTGAGT 1286

```
RESULT 5
CNS0A1K2
LOCUS       CNS0A1K2                1320 bp    mRNA    linear   HTC 06-FEB-2004
DEFINITION  Arabidopsis thaliana Full-length cDNA Complete sequence from clone
            GSLTFB14ZD09 of Flowers and buds of strain col-0 of Arabidopsis
            thaliana (thale cress).
ACCESSION   BX829446
VERSION     BX829446.1  GI:42458363
KEYWORDS    HTC; GSLT; cDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1320)
  AUTHORS   Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
            Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
            Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
  TITLE     Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
            A Combined Approach to Evaluate and Improve Arabidopsis Genome
            Annotation
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1320)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     The sequences are based on single pass reads.
            life Technologies (a division of Invitrogen) members carried out
            full-length libraries construction : Temple G.
            Genoscope members carried out sequencing and annotation : Castelli
            V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
            Schachter V., Weissenbach J., Salanoubat M.
            URGV INRA : Clepet C., Caboche M.
            Annotation is based on the June 2003 version of the Arabidopsis
            genome released by MIPS (Munich Information center for Protein
            Sequences). 5 prime and 3 prime are assembled with Phrap.
            http://www.genoscope.cns.fr/externe/sequences/Banque_EF/Full
            length
            http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES             Location/Qualifiers
     source          1..1320
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="GSLTFB14ZD09"
                     /tissue_type="Flowers and buds"
                     /ecotype="Col-0"
                     /plasmid="pCMVSPORT 6"
                     /clone_lib="GSLT"
     gene            complement(1..1320)
                     /gene="At5g07000"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              1.29e-169
Score:                  1573.00          Length:         1320
Percent Similarity:     92.18%           Matches:        304
Best Local Similarity:  84.92%           Conservative:    26
Query Match:            82.88%           Mismatches:      22
DB:                                      Indels:           6
                                         Gaps:             3

US-10-019-931-3 (1-359) x CNS0A1K2 (1-1320)

Qy    4 SerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHisLysLeuGlu 23
        |||||||||::|||||||||||||||||||||||||||||||||||||:|||||
Db   29 TCAAGCATAAAGAGACATTCCAATGGCGGATCCCAAGTTTCCCATGTGTCACAAGCCCGAG 88
```

```
RESULT 6
LOCUS       CNS09YST          1302 bp    mRNA    linear    HTC 06-FEB-2004
DEFINITION  Arabidopsis thaliana Full-length cDNA Complete sequence from clone
            CNS09YST
ACCESSION   BX831024
VERSION     BX831024.1  GI:42454963
KEYWORDS    HTC; GSLT cDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1302)
  AUTHORS   Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
            Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
            Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
  TITLE     Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
            A Combined Approach to Evaluate and Improve Arabidopsis Genome
            Annotation
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1302)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     The sequences are based on single pass reads.
            Life Technologies (a division of Invitrogen) members carried out
            full-length libraries construction : Temple G.
            Genoscope members carried out sequencing and annotation : Castelli
            V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
            Schachter V., Weissenbach J., Salanoubat M.
            URGV INRA : Clepet C., Caboche M.
            Annotation is based on the June 2003 version of the Arabidopsis
            genome released by MIPS (Munich Information Center for Protein
            Sequences). 5 prime and 3 prime are assembled with Phrap.
            http://www.genoscope.cns.fr/externe/sequences/Banque_EF/Full
            length
            http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES             Location/Qualifiers
     source          1..1302
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="GSLTL502A11"
                     /tissue_type="Adult vegetative tissue"
                     /ecotype="Col-0"
                     /plasmid="pCMVSPORT 6"
     gene            complement(1..1302)
                     /gene="At5g07000"
ORIGIN
```

GSLTL502A11 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).

```
US-10-019-931-3 (1-359) x CNS09YST (1-1302)

Alignment Scores:
Pred. No.:                        1
Score:                      1533.00
Percent Similarity:          91.04%
Best Local Similarity:       82.91%
Query Match:                 80.77%
DB:

                    Length:    1302
                    Matches:    296
                    Conservative: 29
                    Mismatches:  26
                    Indels:       6
                    Gaps:         3
```

Left alignment block (Qy / Db):

```
Qy     24  LeuLeuLeuLysGluGlyGlyLysThrArgAspValProLysAlaGluGluAspGluLeuSer   43
           |||||||||:||||||||::|
Db     89  CTCCCTTAAGGAAGGCAAAAGC-------------GAAGGCCAAGAAGAAGAGGGCTAAGC   136

Qy     44  CysGluPheGlnGluMetLeuAspSerLeuProLysGluArgGlyTrpArgThrArgTyr    63
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    137  TACGAGTTCCAAGAGATGTTGGACTCTCTTCCTAAGGAGAGAGGGAGAAATCGTTAC   196

Qy     64  LeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGlyIleGlnAlaIleIleMetSerPhe   83
           |||||||||||||:|||||||||
Db    197  CTTTACTTATTCCAAGGGTTCCGGTGCCAAGCTAAGGGAATTCAAGCTATCACGTCTTTC   256

Qy     84  GlnLysHisPheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLysSerGly    103
           |||||||||||||||||||||||||||||||||||||||
Db    257  CAAAAACATTTTCAGTCCCTTCCAGACGACGTTGTCCTCGGCCACCATACCTAAAATCTGGC   316

Qy    104  ThrThrThrProLeuLysAlaLeuThrPheThrIleLeuAsnArgHisArgPheAspProVal   123
           ||::|||||||||||||||||:
Db    317  ACAACCTGGTTAAAAGCTTTAACTTTCACCATCCTTACCCCTCATCGGGTTTGATCGGTT   376

Qy    124  ---AlaSerSerThrArgAsnHisProLeuPheThrSerAsnProHisAlaSerLeuValProPhe   142
           ::||||||:|||||||::
Db    377  TCCTCATCAAGTTCCGACCAACCCTCTTCTCACATCCAACCCTCGACCTCGTACCTTTC   436

Qy    143  PheGluThrTyrTyrSerLeuTyrArgAlaAsnGlyLysAspValProAspSerLeuSerPro   162
           ||||||||||:||||:|||||:||
Db    437  TTCGAGTACAAGCTTTACGACAGTTCCCGAATCTCTCGGGGTCTAGCCACAGTCCA   496

Qy    163  ArgThrPheAlaThrHisLeuProPheGlySerLeuGluGluThrIleGlyIleTyrGly    182
           |||||||||||||||||||||||:|
Db    497  AGAACATTCGCAACCCACGTACCGGTCGGTCGCCCTTAAGGATTCCGGTCGGAAATCCCAGT   556

Qy    183  ValLysValAlaTyrIleLeuCysArgArgAsnProPheAspThrPheIleSerSerTrpHisTyr   202
           ||||||||:|||||||||||:
Db    557  GTGAAAGGTTGTCTACCTGTGCCCGGAACCCGTTTGACACATTCATCTCCATGTGGCATTAC   616

Qy    203  ThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaAlaPheAspLeu   222
           ||||:|||||||||::|
Db    617  ATCACACAATCACTTCCGAGTCTTGCTAGACGAAAGCTTTTGATCTA   676

Qy    223  TyrCysArgGly---ValIleGlyPheGlyProPheTrpGluHisMetLeuGlyTyrTyrTrp   241
           ||||||||||||||||||||:||||
Db    677  TATTCCCGGGGATTACTGATCGGATTTGGGCCCGTTTGGGAACACATGTGGGATACTGG   736

Qy    242  ArgGluSerLeuLysArgProGlyLysValPhePheLeuArgTyrGlyIleAspLeuLysLysAsp   261
           ||||||||||:|
Db    737  AGAGAGAGCTTGAAGAGGCCAGAGAAAGTCTTATTTTAAAGTACGAGGATCTCAAAGAA   796

Qy    262  AspIleGluThrArgAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGluGlu   281
           ||||||||:|||||||||||:||:||
Db    797  GACATCGAGACCCAACTTGAACAAGCTAGCAAGTTCTTAGGACTTCCTTTCACCGAAGAA   856

Qy    282  GluGluArgGlyLysValValLysValValAlaIleAlaGluLeuCysSerSerPheGluAsnLeuLys   301
           |||||||||:|||||||||||:||||||
Db    857  GAGGAACAAAGGGGAGTTGTGAAAGCTATCGCTGAGCTTTGTAGCTCTTTTGAGAATCTGAAG   916

Qy    302  LysLeuGluGluLeuValAsnLeuLysSerAsnPheGluAsnArgPheLeuPhe    321
           |||||||||:|
Db    917  AAGTTGGAGGAGTTGGTAAACCTCAAGTCAAATTTGAGAACCGGTTCTTGTTT   976

Qy    322  ArgLysGlyGlyLysSerAspTrpValAsnTyrLeuSerProSerGlnIleValGluArgLeu   341
           ||||||||||||:
Db    977  AGGAAAGGAGGAAGTGAGTGATTGGGTTAACTATTTGTCGCCATCACAAATCGTAGAGAGATTG   1036

Qy    342  SerAlaLeueuValAspAspLysLeuGlyLeuGlySerGlyLeuThrPheArgLeuSer   359
           ||:|||||||||||||||||||||||||||
Db   1037  TCAGCCTTAGTGGATACACAGATTAGCTGGGATACTGTGTCTCACTTTCAGATTGAGT   1090
```

Right alignment block (Qy / Db):

```
Qy     4  SerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHisGlyLeuGlu    23
          |||||:||::||||:|||||||||||:
Db    12  TCAAGCATAAAGAGCATTCCAATGGCGATCCCAAGTTTCTCCAAGTTCATGGGACTAGAG    71

Qy    24  LeuLeuLeuLysGluGlyGlyLysThrArgAspValProLysAlaGluGluAspGluLeuSer    43
          ||||||||||:||||||||::|
Db    72  CTCCTTAAGGAAGGCAAAAGC-------------GAAGGCCAAGAAGAAGAGGGCTAAGC   119

Qy    44  CysGluPheGlnGluMetLeuAspSerLeuProLysGluArgGlyTrpArgThrArgTyr    63
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   120  TACGAGTTCCAAGAGATGTTGGACTCTCTTCCTAAGGAGAGAGGGAGAAATCGTTAC    179

Qy    64  LeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGlyIleGlnAlaIleIleMetSerPhe    83
```

```
RESULT 7
LOCUS        CNS0A0WP               1328 bp    mRNA      linear    HTC   06-FEB-2004
DEFINITION   Arabidopsis thaliana Full-length cDNA Complete sequence from clone
             GSLTPGH36ZF09 of Hormone Treated Callus of strain col-0 of
             Arabidopsis thaliana (thale cress).
ACCESSION    BX831880
VERSION      BX831880.1  GI:42458032
KEYWORDS     HTC; GSLT cDNA.
SOURCE       Arabidopsis thaliana (thale cress)
ORGANISM     Arabidopsis thaliana
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
             rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1  (bases 1 to 1328)
AUTHORS      Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
             Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
             Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE        Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
             A Combined Approach to Evaluate and Improve Arabidopsis Genome
             Annotation
JOURNAL      Unpublished
REFERENCE    2  (bases 1 to 1328)
AUTHORS      Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
             BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT      The sequences are based on single pass reads.
             - Web : www.genoscope.cns.fr)
             Life Technologies (a division of Invitrogen) members carried out
             full-length libraries construction : Temple G.
             Genoscope members carried out sequencing and annotation : Castelli
             V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
             Schachter V., Weissenbach J., Salanoubat M.
             URGV INRA : Clepet C., Caboche M.
             Annotation is based on the June 2003 version of the Arabidopsis
             genome released by MIPS (Munich Information center for Protein
             Sequences). 5 prime and 3 prime are assembled with Phrap.
             http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
             length
             http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
FEATURES             Location/Qualifiers
     source          1..1328
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="GSLTPGH36ZF09"
                     /tissue_type="Hormone Treated Callus"
                     /ecotype="Col-0"
     gene            1..1328
                     /ecotype="Col-0"
                     /plasmid="pCMVSPORT 6"
                     /complement(1..1328)
                     /gene="At5g07010"
ORIGIN
```

US-10-019-931-3 (1-359) x CNS0A0WP (1-1328)

Alignment Scores:		
Pred. No.:		1.97e-164
Score:		1528.00
Percent Similarity:		89.97%
Best Local Similarity:		87.47%
Query Match:		80.51%
DB:		4

Length:	1328
Matches:	1328
Conservative:	314,
Mismatches:	9
Indels:	25
Gaps:	11
	2

```
Qy    8 SerIleProMet-AlaIleProSerPhe-SerMetCysHisHisLeuGlnLeuLeuLysG 27
          |||| |||||   ::::: :|||||  :::  |||| |||| |||| |||
Db   65 AGCAATTCCAATGGCAGGGGGGTCCCAAGTTTCTTTTTTTACAATCTCCACCTCCTTAAAG 124

Qy   27 luGlyLysThr---ArgAspValProLysAlaGluGluGlyLeuSerCysGluP 46
          ||||| ||||        ||||| ||||||||||||| ||||| |||||||
Db  125 AAGGCAAAACTCCCCACCGCTCCCGCAAACCCACAACATGAAGGGCTAAGCTGCCAGT 184

Qy   46 heGlnGluMetLeuAspSerLeuProLysGlyArgGlyTrpAlaArgThrArgTyrLeuTyrI 66
          |||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  185 TCCAAGACATGTTGGATTCTCTCTCCTAAGGAGAGGGGATGGAGAGAACTCGTTACCTTACC 244

Qy   66 euPheGlnGlyPheTrpCysGlnAlaAlaGluIleGlnAlaIleMetSerPheGlnValBH 86
          |||||||||||| |||||||||  :::  :::::::||| ||||||:::::
Db  245 TATTCCAAGGCGGTTTGGTGCCAAGCCAAAAGAGATTCAAGCCATCAATGTCTTTCAAAGGG 304

Qy   86 isPheGlnSerPheLeuGluAsnAspValValLeuAlaThrIleProLeuSerGlyThrThr 106
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  305 ATTTCCAATCCCTCGAGAACGACGTCGTTCTCGCCACCATACCTAAATGGGGTACAACCT 364

Qy  106 rpLeuLysAlaLeuThrPheThrIleLeuLeuAsnArgHisArgPheAspProValAlaSerS 126
          ||||:|||||||||||||||||||||||||||||||:::|||||||||||||||||||||
Db  365 GGCTAAGAGCTTAACTTCACCATCCTTAACCGTCACCGGTTTGATCCGGTTGCTCGA 424
```

```
Qy  104 ThrThrTrpLeuLysAlaLeuThrPheThrIleLeuLeuAsnArgPheAspProVal 123
        |||||||
Db  180 CTTTACTTPATTCCAAGGGGTTCGGTGCCAAGCTAAGGAGGATTCAAGCTATCACGTCTTTC 239

Qy   84 GlnLysHisPheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLySerGly 103
        :::||||::::;|||||
Db  240 CAAAAACATTTCAGTCCCTTCCAGAGAGTTGTTCTCGCCACCATAACTAAATCTGGC 299

Qy  124 ---AlaSerThrAsnHisAlaProLeuPheThrSerAlaAspLeuValProPhe 142
        :::|||||::::;||||||
Db  360 TCCTCATCAAGTTCCGACCACCCCTTCTCAACATCCAACCCGTACCTTTC 419

Qy  143 PheGluTyrLysTyrLysTyrAlaAsnGlyAspValProAspLeuSerGlyLeuAlaSerPro 162
        |||:::;||||||||
Db  420 TTCAGTACTAAGCTTACCCCAACCGAAAAGTTCCGGATCTCTCGGGTCTTAGCCAGTCCA 479

Qy  163 ArgThrPheAlaThrHisLeuProPheGlySerLeuGlyThrIleGlyLysProGly 182
        |||||||||||||||
Db  480 AGAACATTCGCAACCCGTACCGGTTCGGTCGCCCTTAAGGAATTCGGGTCGAGAATCCCAGT 539

Qy  183 ValLysValAlaTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrpHisTyr 202
        |||||||||||
Db  540 TTTPAAGGTTGTTTACCGTGTCCCGGAACCCGTTGACACAATTCATCCATGTTGCATTAC 599

Qy  203 ThrArgAsnAsnAsnIleLysSerGluSerValSerProValAlaThrArgLeuLeuAspAspLeu 222
        ::::||||||||||||
Db  600 ATCAACACACATCACTTCCGAGTCGTCTGCTAGACGAAGCTTTGATCTA 659

Qy  223 TyrCysArgGly---ValIleGlyPheGlyProPheTrpGluHisMetLeuGlyTyrTrp 241
        :::;||||||:::||||||||
Db  660 TATTGCCGGGATTACTGATCGGATTTGGCCCGTTTGGGAACACATGTTGGGATACTGG 719

Qy  242 ArgGluSerLeuLysArgProGlyLysValPheProLeuLysLeuLysAsp 261
        ||||||:::|||||
Db  720 AGAGAGAGCTTTGAAGGCCAGAGAAAGTCTTATTTTAAAGTACGAGGATCTCAAAGAA 779

Qy  262 AspIleGluThrArgAsnLeuLeuArgLeuLeuThrPheLeuGluLeuProPheThrGluGlu 281
        |||::::;|||||||
Db  780 GACANCGAGACCACTTGAAGAGGAGCAAGTTTCTTAGGACTTCCTTTCACCGAAGAA 839

Qy  282 GluGlyArgGlyValValValValAlaIleAlaLeuGluLeuCysSerPheGluAsnLeuLys 301
        |||||::::;|||||
Db  840 GAGGAACAAAAGGAGGGAGTTGTGAAAGCTATCGCTGAAGCTATCGCTGAAGAATCTGAAG 899

Qy  302 LysLeuGluLysValAlaSerAsnLeuLysSerGlyLeuTyrThrPheLeuProPhe 321
        ||||||
Db  900 AAGTTGGAGGTGAACAAGTCAAGCAAAATTGATCCAGAACTATGAGAACCCGGTTCTTGTT 959

Qy  322 ArgLysGlyGlyValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArgLeu 341
        :::||||::::;|||||
Db  960 AGGAAAGGAGGAAGTGAGTCATTTGGTTAACTATCTGCGATCTTCGCCATCACAAGTGGAAGATTG 1019

Qy  342 SerAlaLeuGlnValAlaAspAspGlyLeuGlySerGlyLeuTyrLeuThrPheArgLeu 358
        ||||::::;||||||
Db  1020 TCAGCCTTAGTGGATGGATTGACAAGTTAGCTGGTCTCAACCTTTCACCTTTCAGATT 1070
```

```
RESULT 8
BZ062934/c
LOCUS       BZ062934              757 bp    DNA    linear    GSS 10-OCT-2002
DEFINITION  11c02b11.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
            sequence.
ACCESSION   BZ062934
VERSION     BZ062934.1  GI:23673253
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 757)
  AUTHORS   Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
            Nash,W., Rabinowicz,P.D. and Wilson,R.K.
  TITLE     Whole genome shotgun reads from Brassica oleracea
  JOURNAL   Unpublished (2002)
  COMMENT   Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Plate: 11c02 row: b column: 11
            Seq primer: -21UPpOT forward
```

```
FEATURES             Location/Qualifiers
     source          1..757
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /clone="B.oleracea002"
                     /db_xref="taxon:3712"
                     /clone_lib="B.oleracea002"
                     /note="Vector: pOTw13; Whole genome shotgun library from
                     flowering buds. DNA was purified from a crude nuclear
                     prep using Brassica oleracea TO1000DH3 buds provided by
                     Thomas Osborn at the University of Wisconsin. Genomic
                     DNA was provided by Pablo Rabinowicz (CSHL) and the
                     shotgun library prepared at Washington University Genome
                     Sequencing Center."
ORIGIN
```

```
Class: shotgun
High quality sequence start: 52
High quality sequence stop: 551.
```

Alignment Scores:

```
Pred. No.:                              6.39e-116
Score:                                  1103.50
Percent Similarity:    92.53%    Length:              757
Best Local Similarity: 83.40%    Matches:             201
Query Match:           58.14%    Conservative:         22
                                 Mismatches:           15
                                 Indels:                3
DB:                              Gaps:                  1
```

US-10-019-931-3 (1-359) x BZ062934 (1-757)

```
Qy    126 erThrAsnHisProLeuPheThrSerAsnProHisAspLeuValProPheGlyTyrL 146
Db    425 GTACCAACCACCCTCTTTCACTTCAAACCCTCATGACCTTGTACCTTTCTTCGAGTACA 484

Qy    146 ysLeuLeuTyrAlaAsnGlyAspValProAspLeuSerGlyLeuValaSerProArgThrPheA 166
Db    485 AGCTTTACGCCAACGGAGATGTTCCCGATCTCCGGGTCTAGCCCAGTGCAAGAACGTTCG 544

Qy    166 laThrHisLeuProPheGlySerLeuValGluThrIleGluGlyValValLysValV 186
Db    545 CAACCCCACTTACCGTTCGGTTCCCTAAAGGAAGATCGAGAAAACCCGGTGTGAAGGTCG 604

Qy    186 alTyrLeuCysArgArgAsnProPheAspThrPheIleSerSerTrpHisTyrThrAsnAsnI 206
Db    605 TGTACTTGTGCCGGAACCCGTTCAAACATTACTCTCTGGGCATTACCACACAACACA 664

Qy    206 leLysSerGluSerValSerProValLeuLeuAspGlnIleAlaPheAspLeuThrTyrCysArg 226
Db    665 TCAAATCCGAGTGAGTGAGTGACCAAGTCTTGATCTCGTTATTCGCCGGG 724

Qy    226 lyValValIleGlyPheGlyProPheTrpGluHisMetLeuGlyTyrThrAspGluSerLeuL 246
Db    725 GAGTGGATCCGGGTTTGGGCCCGTTTGGGAACACATGTTGGACATACTGGAGACGAGCTTGA 784

Qy    246 ysArgArgProGluLysValPheThrLeuArgTyrGluAspAspAspIleGluThrA 266
Db    785 AGAGACCAGAGAAAAGTCTTCTCTTTAAGGTACGACGAGCCATCGACACCA 844

Qy    266 snLeuLysArgGlnLeuAlaThrPheLeuGluProPheThrGluLysGluGluGlyGlyG 286
Db    845 ACTTGAAGAGCGGCTTGCAACTTTCTTCACCGAGAAAGAAGAAGGAGAAGGAGGAGAGA 904

Qy    286 lyValValLysAlaIleAlaGluLeuCysSerPheGluAsn---LeuLysGlyLeuGluGluVa 305
Db    905 GAGTTGTTAAGGCTATCGCCGAGCTCTGGCTTCGGTGCCAAGGCCAAGGAGATT 964

Qy    305 l-AsnLysSerAsnLysSerIleLeuLysAsnPheGluLeu---PheAlaArgLysGl 324
Db    965 CGAACAAGTGAAACAAGTGGATCAAGAACTTTGAGAATCAATCCTTGTTCGGAAAGG 1024

Qy    324 yGluValSerAspTrpValAsnTyrLeu---SerProSerGlnValGluArgLeuSerAla 343
Db   1025 AGAAGTGAGTGAGTTGGGTTCACTATTTGTGCACCCTTCACAAGTGGAAAAGAATTGTCAGCC 1084

Qy    344 LeuValAlaAspAsp---LysLeuGlyGlySerGlyLeuThr 355
Db   1085 TTAGTGGATGACCACGCCTAGGTGGATCGTTCTCACT 1123
```

```
Qy     38 GluAspGluGlyTyrLeuSerCysGluPheGlnGluMetLeuAspSerLeuProLysGluArg 57
Db    755 GAAGAGAAGGGCTAAGCTACGAGTTCTTAGAGATGTTGGACTCTCTTCCTAAGGAGAGA 696

Qy     58 GlyTyrTrpArgThrArgTyrLeuTyrLeuPheGlnGlyGlyPheGlnAlaGlyGluIle 77
Db    695 GGTTGGAGAACTCGTCATCTTTACCTGTTCCAAGGCTTTGGTGCCAAGGCAGGAGATT 636

Qy     78 GlnAlaIleMetSerProPheGlnLysHisPheGlnSerLeuGluAsnAspValValLeuAla 97
Db    635 CATGCCAATCAYGTCCTTCCAGAACACATTTCAAGGCCTTGGTGCCAAGGACGACGGAATT 576

Qy     98 ThrIleProLysSerGlyThrThrTrpLeuTyrPheThrIleLeuAsnArg 117
Db    575 ACCATACCAAAATCCGGTACAACACATGGTTAAAAGCTTAAACATTTACCCTCTTAACCGA 516

Qy    118 HisArgPheAspProValAlaSerSerThrAsnHisProLeuProPheThrSerAsnProHis 137
Db    515 CACCGGTTTGATCCGGTT-------TCTGACCACCCTCTCCATCAAACCCTCAT 465

Qy    138 AspLeuValProPhePheGluTyrLeuTyrLeuAlaAsnGlyAspValProAspLeuSer 157
Db    464 GACCTCGTACCTTTCTTCGAGTACAAGCTTTACGCCAACGGAGAAGTTCCCGATCTCCGG 405

Qy    158 GlyLeuValaSerProArgThrPheAlaThrHisLeuProPheGlySerLeuValGluThr 177
Db    404 AGTCTAGCCTAGTCCAAGAACATTCGACAACACACGTCCCTTCGGTTCCTCCAAGGGCTCA 345

Qy    178 IleGluGlyValLysValProGlyValValLysValTyrLeuCysArgArgAsnProPhe 197
Db    344 ATCGAGGAAACCCGGAGTGAAGGTTGTGTTACTTGTGCCGGTTCAGAACGTTCATC 285

Qy    198 SerSerTrpHisTyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAsp 217
Db    284 TCCTCGTTCGGCATTTCAGCACAACAACAGCATAAAAGTCGGAGTGAGTGAGTGACCAG 225

Qy    218 GlnIleAlaPheAspLeuThrTyrCysArgGlyValIleGlyPheGlyProPheTrpGluHisMet 237
Db    224 GAAGGGTTTGATCTGTATTCGAGGGAGAGTGATCGGCGTTTGGGAACACATG 165

Qy    238 LeuGlyTyrThrAspGluSerLeuLysArgArgProGluLysValPheThrLeuArgTyrGlu 257
Db    164 CTGGGGGTACTCGAGACGAGCTTGAGAGACCAGAGGAAGTCTTGTTCTTTAAAGTACGA 105

Qy    258 AspLeuLysAspAspAspIleGluThrAsnLeuLysArgGlnLeuAlaThrPheLeuGluLeuPro 277
```

RESULT 9

BH733391

```
LOCUS       BH733391    811 bp    DNA    linear   GSS 20-FEB-2002
DEFINITION  BOMIU60TR BO_2_3_KB Brassica oleracea genomic survey sequence.
ACCESSION   BH733391
VERSION     BH733391.1  GI:18838786
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 811)
  AUTHORS   Ayele,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
            Utterback,T.R., Wortman,J.R., White,O.R. and Town,C.D.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea and its
            application to gene discovery and annotation in Arabidopsis
  JOURNAL   Genome Res. 15 (4), 487-495 (2005)
   PUBMED   15805490
COMMENT     Other GSSs: BOMIU60TF
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            Contact: Chris Town
            Seq primer: TR
            Class: sheared ends.
            DNA is from a doubled haploid provided by Tom Osborn.
FEATURES             Location/Qualifiers
     source          1..811
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOMIU60"
                     /clone_lib="BO_2_3_KB"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
ORIGIN
```

Alignment Scores:
```
Pred. No.:               9
Score:             8.79e-115
Percent Similarity:   93.72%
Best Local Similarity: 83.26%
Query Match:          57.64%
DB:
```

us-10-019-931-3 (1-359) x BH733391 (1-811)

```
Length:      811
Matches:     199
Conservative: 25
Mismatches:   15
Indels:        0
Gaps:          0
```

```
Qy  119  ArgPheAspAlaProValAlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAsp  138
         |||||||||||  ::   :|||||||||||||||||||||||
Db    3  CGGGTTTGATCCGGTTTCTTCAACCACCGACCACTCCTCAACATCAAAACCCTCATGAC   62

Qy  139  LeuValProPheGluTyrArgLeuTyrAlaAsnGlyAspValProSerGluSerGly     158
         ||||||||||||||||||||||||||||||||||||||||
Db   63  CTCGTACCTTTCTTCGAGTACAGGCTTTACGAGACGGTGATGTTCCTGAGTCTCGGT    122

Qy  159  LeuAlaSerProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIle  178
         ||||||||||||||||||||||:|||::::::|||
Db  123  TTAGCCAGTCCGAGAACATTCGCAACCCACGTCCCTTTCGGTTCCCTCAAGGACATC   182

Qy  179  GluLysProGlyValValValTyrLeuCysArgAsnProPheAspThrPheLeuSer    198
         |||::||||||||||||||||||||
Db  104  GATCTGAAAGAAGATATTGAGACTAACTTGAAGAAGCTTGCAAGTTTCTTGGGGGTTTCT  45
```

RESULT 10

CD837475

```
LOCUS       CD837475    727 bp    mRNA    linear   EST 10-JUL-2003
DEFINITION  BN45.052G01F020104 BN45 Brassica napus cDNA clone BN45052G01, mRNA
            sequence.
ACCESSION   CD837475
VERSION     CD837475.1  GI:32519415
KEYWORDS    EST.
SOURCE      Brassica napus (rape)
  ORGANISM  Brassica napus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 727)
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1..727
                     /organism="Brassica napus"
                     /mol_type="mRNA"
                     /cultivar="Jet Neuf"
                     /db_xref="taxon:3708"
                     /clone="BN45052G01"
                     /tissue_type="seed"
                     /clone_lib="BN45"
ORIGIN
```

Alignment Scores:
```
Pred. No.:
Score:          2.4e-100
Length:              727
Matches:             186
```

```
Db  183  GAGGAACCAGGACGTGAAAGGTGGGTTGGTGTACTTTTGCCGAAACCCGTTCGGACACATTCATCTCC  242

Db  199  SerTrpHisTyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGln  218
         |||||||||
Db  243  TCGTGGCATTACACTCAACAACATAAAGTCGGAGAGTCGGTCCGTCCGTTAGGAAGA     302

Qy  219  AlaPheAspAsnLeuTyrCysArgGlyValIleIleGlyPheGlyProPheTrpGluHisMetLeu  238
         ||||||||||||||||||||||||
Db  303  GGGTTTGATCCTACTGCCCGAGGGGGTGATCGGGTTCGGACCGTTTGGGAACAACATGTTG       362

Qy  239  GlyTyrTrpArgGluSerLeuLeuLysArgProGlyLysLysValPhePheLeuArgTyrGluAsp  258
         ||||||||||||||||||||||::::
Db  363  GGATACTGGAGAAAGAGCTTGAGCAACACCAGAGAAAAGTCTTGTTCTTACGTTACGAGGAT        422

Qy  259  LeuLysAspAspAspIleGluThrPheAsnLeuArgLeuHisLeuAlaThrPheLeuGluLeuProPhe  278
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  423  CTGAAACAAGACAATGGAGTCTAACTTGAGACACCAGAGAAGTTCTTGGACGGTTCCTTTC        482

Qy  279  ThrGluGluGluGluArgGlyValValValLysValAlaLysGluAlaGluAlaGluLeuCysSerPheGlu  298
         ::::||||||||||||||||||||:::||||||||||||||||
Db  483  ACGGAAGAAGAGGAACAAAAAGGTGTGGTGAAGGTCGCCAAGGAAGCTTGCAGCTTCGAG         542

Qy  299  AsnLeuLysGlyLeuGluValValAsnLysSerIleIleLysAsnPheGluAsnArg  318
         ||||||||||||:::|||||||:||||||||||||||
Db  543  AATCTAAAGGAAGCTGAGGAGGTGAACAAGTCAAGCAAATCGATCAAGAACTTTGAGAATAGG      602

Qy  319  PheLeuPheArgLysGlyGlyValSerAspTrpValAsnTyrLeuSerProSerGlnVal  338
         ||||||||||||||||||||||||||||||||||||||||
Db  603  CACTTCTTCAGGAAAGGAGGAGTGAGTGATTGGGTTAACTATCTCTCGCCACCACAGGCC       662

Qy  339  GluArgLeuSerAlaLeuValLeuAspAspArgLysGlyLeuGlySerGlyLeuLeuThrPheArg  357
         |||:::|||||||:||||||||||||||
Db  663  GAAAGATTGTCAGCCTTAGTCGAATGACAGTTAGGTGGTTATGGGTCTCACTTTCAGG         719
```

Percent Similarity: 90.39%   Conservative: 21
Best Local Similarity: 81.22%   Mismatches: 14
Query Match: 51.00%   Indels: 8
DB: 6   Gaps: 5

us-10-019-931-3 (1-359) x CD837475 (1-727)

```
Qy   1 MetAlaThrSerSerMetLysSerIlePro---MetAlaIleProSerPheSerMetCys 19
       ||||||||||||||||||||||||||||:::|  ||||||||||||||||||||||||||
Db  51 ATGGCGACCTCAAGCATCAAGAGTGTTCCAATAATGGCGATCCCAAGTTTCTCCATTTGC 110

Qy  20 HisLysLeuGluLeuLeuLys---GluGlyLysThrArgArgAspValProLys-----Ala 36
       ||||||||||||||||||||    |||||||||||||:::||||   |||||||||    |||
Db 111 CACAAGCACGAGCTCCTCAAGGAAGGAAAAGGCCAAGAC---CCGAAGCGCCCAAGAA 167

Qy  37 GluGluAspGluGlyLysLeuSerCysGluPheGlnGlyMetLeuAspSerLeuProLysGlu 56
       ||||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db 168 GAAGAAGAAGGGCTAAGCTACGAGTTCTTAGAGATGTTGGACTCTCTTCCTAAGGAG 227

Qy  57 ArgGlyTrpArgThrArgTyrLeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGlu 76
       |||||||||||:::||||||||||||||||||:::||||||||||||||||||||||
Db 228 AGAGGTTGGAGAACTCGTCAATCTTTACCTGTTCCAAGGCTTTTGGTGCCAAGGAG 287

Qy  77 IleGlnAlaIleMetSerPheGlnIleHisProLeuPheThrSerAsnProLysAspValIleLeu 96
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 288 ATTCAATGCCATCATGTCCTTCCAGAAAACATTTCAAGACTCTCCAAAAGATGTTATTCTC 347

Qy  97 AlaThrIleProLysSerGlyThrThrPheHisLeuProPheGlySerLeuLysAsn 116
       ||||||||||||||||||||||::::::|||||||||||||||||||:::|||||
Db 348 GCCACCATACCAAAATCCGGTACAACATGGTTAAAAGCTTTACCCTTCCTTCCTTAAC 407

Qy 117 ArgHisArgPheAspProValAlaSerSerThrArgAsnHisProLeuPheThrSerAsnPro 136
       ||||||||||||||||||||--------:::|||||:::||||||||||||||||||
Db 408 CGACACCGGTTTGATCCGGTT--------TCTGACCACCCTCTCTCCACATCAAACCCT 458

Qy 137 HisAspLeuValProPhePheGluThrTyrLysAlaLeuAlaAsnGlyAspValProLeuAsn 156
       ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||
Db 459 CATGACCTCGTACCTTTCTTGGAGTACAAAGCTTAACGGAGAAGTTCCTGATCTC 518

Qy 157 SerGlyLeuAlaSerProArgThrArgPheAlaThrHisLeuProPheGlySerLeuLysGlu 176
       |||||||||||||||||||||||::|||||||||:::||||||||||||||||||||
Db 519 TCCGGTCTAGCTAGTCCAAGAACAATTCGGCAACACACGCTTCCCTTCGGTTCTCTCAAGGAG 578

Qy 177 ThrIleGluLysProGlyValLysValValTyrLeuCysArgGlyAsnProPheAspThrPhe 196
       ||||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db 579 TCCATCGAGAAGAACCCGGAGCGAAGGTTGTGTACTTGTGTCGGACAACCCGGTTCGGAGGTTC 638

Qy 197 IleSerSerTrpHisThrArgAsnIleLysSerGluSerValSerProValLeuLeu 216
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 639 ATCTCCTCGTGGCATTTCAGCAACACAGCATAAAGTCGGAGTCGGTGACTCTCGGTTG 698

Qy 217 AspGlnAlaPheAspLeuTyrCysArg 225
       :::::
Db 699 GAAGAAGAGGGGTTGATCTGTATTGCAGG 725
```

RESULT 11

LOCUS   CD834047   734 bp   mRNA   linear   EST 10-JUL-2003
DEFINITION BN45_040G17F011018 BN45 Brassica napus cDNA clone BN45040G17, mRNA
            sequence.
ACCESSION   CD834047
VERSION   CD834047.1 GI:32515987
KEYWORDS   EST.
SOURCE   Brassica napus (rape)
  ORGANISM Brassica napus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1 (bases 1 to 734)
  AUTHORS   Genoplante.
  TITLE   Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Genoplante

            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).

FEATURES   Location/Qualifiers
     source   1..734
                     /organism="Brassica napus"
                     /mol_type="mRNA"
                     /cultivar="Jet Neuf"
                     /db_xref="taxon:3708"
                     /clone="BN45040G17"
                     /tissue_type="seed"
                     /clone_lib="BN45"
ORIGIN

Alignment Scores:
Pred. No.:
Score: 1.03e-96
                           936.50

Percent Similarity: 89.22%   Matches: 188
Best Local Similarity: 81.03%   Conservative: 19
Query Match: 49.34%   Mismatches: 17
DB: 6   Indels: 10
                                         Gaps: 5
Length: 734

US-10-019-931-3 (1-359) x CD834047 (1-734)

```
Qy   1 MetAlaThrSerSerMetLysSerIlePro---MetAlaIleProSerPheSerMetCys 19
       ||||||||||||||||||||||||||||:::|  ||||||||||||||||||||||||||
Db  53 ATGGCGACCTCAAGCATCAAGAGTGTTCCAATAATGGCGATCCCAAGTTTCTCCATTTGC 112

Qy  20 HisLysLeuGluLeuLeuLys---GluGlyLysThrArgArgAspValProLys-----Ala 36
       ||||||||||||||||||||    ||||||||||||:::||||   ||||||||||    |||
Db 113 CACAAGCACGAGCTCCTCAAGGAAGGAAAAGGCCAAGAC---CCGAAGCGCCCAAGAA 169

Qy  37 GluGluAspGluGlyLysLeuSerCysGluPheGlnGlyMetLeuAspSerLeuProLysGlu 56
       ||||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db 170 GAAGAAGAAGGGCTAAGCTACGAGTTCTTAGAGATGTTGGACTCTCTTCCTAAGGAG 229

Qy  57 ArgGlyTrpArgThrArgTyrLeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGlu 76
       |||||||||||:::||||||||||||||||||:::||||||||||||||||||||||
Db 230 AGAGGTTGGAGAACTCGTCAATCTTTACCTGTTCCAAGGCTTTTGGTGCCAAGGAG 289

Qy  77 IleGlnAlaIleMetSerPheGlnIleHisProLeuPheThrSerAsnProLysAspValIleLeu 96
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 290 ATTCAATGCCATCATGTCCTTCCAGAAAACATTTCAAGACTCTCCAAAAGATGTTATTCTC 349

Qy  97 AlaThrIleProLysSerGlyThrThrPheHisLeuProPheGlySerLeuLysAsn 116
       ||||||||||||||||||||||::::::|||||||||||||||||||:::|||||
Db 350 GCCACCATACCAAAATCCGGTACAACATGGTTAAAAGCTTTACCCTTCCTTCCTTAAC 409

Qy 117 ArgHisArgPheAspProValAlaSerSerThrArgAsnHisProLeuPheThrSerAsnPro 136
       ||||||||||||||||||||----------:::|||||:::||||||||||||||||
Db 410 CGACACCGGTTTGATCCGGTT----------TCTGACCACCCTCTCTCCACATCAAACCCT 460

Qy 137 HisAspLeuValProPhePheGluThrTyrLysAlaLeuAlaAsnGlyAspValProLeuAsn 156
       ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||
Db 461 CATGACCTCGTACCTTTCTTGGAGTACAAAGCTTAACGGAGAAGTTCCTGATCTC 520

Qy 157 SerGlyLeuAlaSerProArgThrArgPheAlaThrHisLeuProPheGlySerLeuLysGlu 176
       |||||||||||||||||||||||::|||||||||:::||||||||||||||||||||
Db 521 TCCGGTCTAGCTAGTCCAAGAACAATTCGGCAACACACGCTTCCCTTCGGTTCTCTCAAGGAG 580

Qy 177 ThrIleGluLysProGlyValLysValValTyrLeuCysArgGlyAsnProPheAspThrPhe 196
       ||||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db 581 TCCATCGAGAAGAACCCGGAGCGAAGGTTGTGTACTTGTGTCGGACAACCCGGTTCGGAGGTTC 640

Qy 197 IleSerSerTrpHisThrArgAsnIleLysSerGluSerValSerProValLeuLeu 216
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 641 ATCTCCTCGTGGCATTTCAGCAACACAGCATAAAGTCGGAGTCGGTGACTCA-GTCTCGTTG 699

Qy 217 AspGlnAlaPheAspLeuTyrCysArgGlyValIle 228
```

RESULT 12

```
Db          AU238005

              |||||||||||||||||
Qy    700 GAAGAAGGGTT-GATCTGTATTGCAGGGAGTGATA 734
              |||||||||||||||||
Db          :::::
```

```
LOCUS           AU238005                   671 bp    mRNA          linear
DEFINITION      AU238005 RAFL16 Arabidopsis thaliana cDNA clone RAFL16-75-F02 5',
                mRNA sequence.
ACCESSION       AU238005
VERSION         AU238005.1  GI:1987174
KEYWORDS        EST.
SOURCE          Arabidopsis thaliana (thale cress)
  ORGANISM      Arabidopsis thaliana
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
                rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE       1  (bases 1 to 671)
  AUTHORS       Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
                Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J.,
                Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A.,
                Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
  TITLE         Large scale analysis of Arabidopsis full-length cDNA
  JOURNAL       Unpublished (2002)
COMMENT         Contact: Motoaki Seki
                Plant Functional Genomics Research Group
                RIKEN Genomic Sciences Center
                3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
                Tel: 81-298-36-9359
                Fax: 81-298-36-9060
                Email: mseki@rtc.riken.go.jp
                An Arabidopsis full-length cDNA library was constructed essentially
                as reported previously (Seki et al., 1998).cDNA cleaved with BamHI
                and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
                al., submitted for publication) digested with BamHI and SalI. This
                clone is in a modified pBluescript vector. Please visit our web
                site (http://www.gsc.riken.go.jp/e/plant/index _e.html) for further
                details.
FEATURES             Location/Qualifiers
     source          1..671
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="RAFL16-75-F02"
                     /lab_host="DH10B"
                     /clone_lib="RAFL16"
                     /note="Site_1: BamHI; Site_2: SalI; dark-grown"
ORIGIN
```

US-10-019-931-3 (1-359) x AU238005 (1-671)

Alignment Scores:
Pred. No.:   1.8le-94
Score: 916.50
Percent Similarity: 90.14%
Best Local Similarity: 83.10%
Query Match: 1

Length: 671
Matches: 177
Conservative: 15
Mismatches: 16
Indels: 5
Gaps: 2

```
Qy      4 SerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHisLysLeuGlu 23
            |||||||||::::|||||||||||||::|||||||||||||
Db     45 TCAAGCATAAAGAGCAATTCCAATGGCGGATCCCAAGTTTCTCCATGTGTCACAAGCTGAG 104

Qy     24 LeuLeuLysGlyLysGlyThrArgAspValProLysGluAspGlyLeuSer 43
            |||||||||||||||||::
Db    105 CTCCTTAAGGAAGGCAAAAGC---------GAAGGCCAAGAGAAGAAGGGGTTGAGC 152

Qy     44 CysGluPheGlnGluMetLeuAspSerLeuProLysGluArgGlyTrpArgThrArgTyr 63
            ||||||||||||||||||||||||::|||||||||||||||
Db    153 TACGAGTTCCAAGAGATGTTGGACTCTCTTCCTAAGGAGAGAGGAGGGACGACGTATTAC 212

Qy     64 LeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGlnIleGlnAlaIleMetSerPhe 83
            ||||||||||||::::|||||||||||||||||||||||||||||
Db    213 CTTTACTTATTCCAAGGGTTTCGGTGCCAAGCTAAGGAGAATTCAAGCTATCACGTCTTTC 272
```

RESULT 13

```
Db          CV130872
```

```
LOCUS           CV130872                   875 bp    mRNA          linear
DEFINITION      X9SP07a08 Populus stem seasonal library Populus deltoides cDNA,
                mRNA sequence.
ACCESSION       CV130872
VERSION         CV130872.1  GI:51874792
KEYWORDS        EST.
SOURCE          Populus deltoides
  ORGANISM      Populus deltoides
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
                rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE       1  (bases 1 to 875)
  AUTHORS       Park,S. and Han,K.-H.
  TITLE         Gene expression profile during seasonal growth cycle in poplar tree
  JOURNAL       Unpublished (2003)
COMMENT         Contact: Kyung-Hwan Han
                Department of Forestry
                Michigan State University
                126 Natural Resources, East Lansing, MI 48824-1222, USA
                Tel: 517 353 4751
                Fax: 517 432 1143
                Email: hanky@msu.edu.
FEATURES             Location/Qualifiers
     source          1..875
                     /organism="Populus deltoides"
                     /mol_type="mRNA"
                     /strain="ILL-129"
                     /db_xref="taxon:3696"
                     /tissue_type="stem"
                     /dev_stage="1 year old"
                     /clone_lib="Populus stem seasonal library"
ORIGIN
```

US-10-019-931-3 (1-359) x CV130872 (1-875)

Alignment Scores:
Pred. No.:   1.26e-92
Score: 902.00
Percent Similarity: 75.62%
Best Local Similarity: 56.54%
Query Match: 7

Length: 875
Matches: 160
Conservative: 54
Mismatches: 65
Indels: 4
Gaps: 1

```
Qy     84 GlnLysHisPheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLysSerGly 103
            :::||||||||||||::||||||||||||||||||||||||
Db    273 CAAAAACATTTTCAGTCCCTTCAGACAGCAGACGTTGTCCTGCCACAATACCTTAAATCTGGC 332

Qy    104 ThrThrTrpLeuLysAlaLeuThrPheThrIleLeuAsnArgHisArgPheAspProVal 123
            ||||||||||||||::||||||||||||||:::||||||||||||||
Db    333 ACAACCTGGTTAAAAGCTTTAACTTCACCATCTTAAACCGTAACGTGTCATCGGTTGATCCGGTT 392

Qy    124 ---AlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAspLeuValProPhe 142
               :::|||||||::||||||||||||:::||||||||
Db    393 TCCTCATCAAGTTCCGACCACCCTCTTCACATCCAACCCTCACGACTCGTACCTTTC 452

Qy    143 PheGluTyrGlyLeuTyrArgAlaAsnGlyAspValProAspLeuSerGlyLeuAlaSerPro 162
            |||||||||||||||::||||||||||||||||||||||||
Db    453 TTCGAGTACGCAAGCTTTACGCCAACGGAAAATGTTCCGGATCTCTCAGGCCTAGCCAGTCCA 512

Qy    163 ArgThrPheAlaThrHisLeuProPheGlySerLeuLeuLysGluThrIleGluGlyProGly 182
            :::||||||||||||:::|||||||||||||||::|||||||||||||
Db    513 AGAACATTCGCAACCCACGTACCGTTCGGTCGCCCTTAAGGATTCGGAGAATCCCAGT 572

Qy    183 ValLysValValTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrpHisTyr 202
            ||||||||||||||||||||||||||||||||||||||
Db    573 GTGAAGGTTGTTTACCTGTGCCGGAACCCGTTTGACACATTCATCTCCATGTGGCATTAC 632

Qy    203 ThrAsnAsnIleLeuSerGluSerValSerProValLeu 215
            |||||||||||||||||||||
Db    633 ATCAACAACAACATCACTTCCGAGTTCAGTGAGCGCAGTCTTG 671
```

```
Qy   72  CysGlnAlaLysGlyGlnGluIleLeuGlnAlaIleMetSerPheGlnGlySerLeuGlu   91
             ::|:|||||::|:|:|||
Db   38  TGTCCTGAAATGCAGATTCCTGCTGTGATTCCATTTCAAAGCACTTCCTAGCACAGAAA    97

Qy   92  AsnAspAlaValLeuAlaThrIleProLysSerGlyThrThrTrpLeuLysSerAlaLeuThr  111
             :::
Db   98  ACAGATACTATACTACTCACCATGCCTAAAACCAGGCACAACATGGTTGAAAGCCTTGGCA  157

Qy  112  PheThrIleLeuAsnArgHisArgPheAspProValAlaLaSerThrAsnHisProLeu   131
             |:::|||:|||
Db  158  TTTTCCATTATTGAAATCGTGCAAAATATACACCCTCCGTGCAGCC---------CCCTTG  205

Qy  132  PheThrSerAsnProHisAspLeuValProPheGluTyrTyrLysLeuTyrAlaAsnGly  151
             :::|||
Db  206  AACTCTGTCAACCCTCATGATCTTGTACCTTTCTTTGAGTTTTACAAATTAAC  265

Qy  152  AspValProAspLeuSerGlyLeuValAlaSerProArgThrPheAlaThrHisLeuProPhe  171
             ::|||||:|||
Db  266  CAACTTCCTGACCTGTCTACCTTTCCATCCCCTAGAATGTTTGCTACTCATGTGCCCATAT  325

Qy  172  GlySerLeuGlyThrIleGluGlyLysProGlyValValValValTyrLeuCysArgAsn  191
             |||:|:::
Db  326  CCATCACTACCGGATTCCATCAAAGAACTCCGGCGTTGTTTATCTTTGCACGGAAT  385

Qy  192  ProPheAspThrPheIleSerSerThrHisTyrThrAsnAsnIleLeuLysSerGluSerVal  211
             |||:||||:::
Db  386  CCCTTTTTGACCACCGTATTAGGGTATTGGAGAGAAAGCTTAGAGAGAAGAGAGGTT  445

Qy  212  SerProValLeuLeuLeuAspPheAlaGlnAlaPheAspLeuPheGlyPheGly   231
             |||:||:||:||:||
Db  446  GGGCCATTATCTTTGGAGGATTGTTCGATAGGTTTGCAATGGACATTGGAGGATTCGGT  505

Qy  232  ProThrTrpGluHisMetLeuGlyTyrTrpArgGluSerLeuLysSerArgProGluValVal  251
             :::|||||
Db  506  CCCTTTTTGACCACCGTATTAGGGTATTGGAGAGAAAGCTTAGAGAGAAGAGAGGTT  565

Qy  252  PhePheLeuArgTyrGluAspLeuLysSerArgIleLeuThrArgLeuAla   271
             |::|||||
Db  566  CTGTTTCTCACGTATGAGGACATGAAAAGGACACATTAATTCTCAGATGAATGGGAAGCT  625

Qy  272  ThrPheLeuGluLeuProThrThrGluGluGluGluGluGluGluLysLysSerAsnLysSerSer  291
             |||||::|:||
Db  626  GAGTTCCTGGGCGTGTCCTTTTTTCCTTTGGAGAACAGCAAGATGGGTTGTGTGAAGAATA  685

Qy  292  AlaGluLeuCysSerPheGluLeuAsnArgPheLeuArgGlyGlyValSerAspTrpValAsn  311
             |||:::
Db  686  TCAAAGTTGTGTGTAGCTTCAGCAATTTGAAGGACAACAACAGAGACTGGGCAAAGTCT  745

Qy  312  IleLeuLysAsnArgPheLeuGlyLysGlyGluValSerAspTrpValAsn   331
             |::::|||
Db  746  ATCCCACACTATGAAGAACAAGACTCTCTTAGAAGAGGTGAAGTGGGGGGAATTGGGTCAAT  805

Qy  332  TyrLeuSerProSerGlnValGluArgLeuSerAlaLeuValAspAspLysLeuGlyLeuGly  351
             ::|||||
Db  806  TACCTTACTCCTGAGANGGTGGATCGTTGAACAAAATCACCGAACAAAAAGCTGGCTGGT  865

Qy  352  SerGlyLeu  354
             ||||||||
Db  866  TCTGGTTTG  874
```

RESULT 14

LOCUS   CX043920   817 bp   mRNA   linear   EST 09-DEC-2004
DEFINITION UCRCS07_14C09.g Parent Washington Navel Orange Thrip-Challenged
            Flavedo cDNA Library UCRCS07 Citrus sinensis cDNA clone
            UCRCS07-14C09-F18-1-7.g, mRNA sequence.
ACCESSION   CX043920
VERSION   CX043920.1 GI:56528211
KEYWORDS   EST.
SOURCE   Citrus sinensis
  ORGANISM Citrus sinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids II; Sapindales; Rutaceae; Citrus.
REFERENCE   1 (bases 1 to 817)
  AUTHORS   Close,T.J., Roose,M.L., Federici,C.F., Fenton,R.D., Watkins,P.,
            Morse,J., Wanamaker,S., Lyon,M., Ye,X.R., Jang,C., Quinitio,C.,
            Ikeda,J., Collin,M., Kacar,Y., Landry,B., Hubert,N., Laforest,M.,
            Landry,J. and Ligonde,A.
  TITLE   Development of EST Resources and New Genetic Markers for California
            Citrus - Parent Washington Navel Orange Thrip-Challenged Flavedo
            cDNA Library UCRCS07
  JOURNAL   Unpublished (2004)
COMMENT   Contact: Timothy J. Close
            Department of Botany & Plant Sciences
            University of California
            Riverside, CA 92521-0124, USA
            Tel: 909-787-3318
            Fax: 909-787-4437
            Email: timothy.close@ucr.edu
            Seq primer: T3.
FEATURES   Location/Qualifiers
     source   1..817
                     /organism="Citrus sinensis"
                     /mol_type="mRNA"
                     /cultivar="Parent Washington Navel"
                     /db_xref="taxon:2711"
                     /clone="UCRCS07-14C09-F18-1-7.g"
                     /tissue_type="flavedo"
                     /dev_stage="11 year old trees"
                     /lab_host="E. coli TJC121"
                     /clone_lib="Parent Washington Navel Orange
                     Thrip-Challenged Flavedo cDNA Library UCRCS07"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site 1: EcoRI; Site 2: XhoI; Trees were grown in the field
                     at University of California, Riverside using standard
                     horticultural practices. Young fruits were placed in a
                     cage with thrips (Scirtothrips citri). Infestations were
                     conducted by Watkins (Morse lab) and flavedo collected by
                     Federici (Roose lab) from May-June 2003. The thrips were
                     collected from naturally infested Rhus plants by sucking
                     into a tube. A flexible hose was attached to a tube that
                     extended into a covered vial, and another tube stuck out
                     of the vial at a right angle. The bent tube was held above
                     the thrip, sucking on the flexible tube created a vacuum,
                     pulling the thrips into the vial. Thrips were knocked off
                     the Rhus plant onto a manila folder, then only second
                     instars were captured. This was done repeatedly until
                     enough were obtained. Approximately 7-10 thrips were caged
                     on each fruit within a plastic vial made of a 8 cm long by
                     5 cm diameter tube that had a very fine mesh organdy
                     fabric glued to the bottom. The plastic cap was slit from
                     the edge to the center so it could be slipped over the
                     stem of the fruit. It was put in place, the thrips were
                     knocked into the vial and it was fastened onto the cap
                     then all gaps were closed with masking tape. The thrips
                     naturally move up to the fruit. The cages were left in
                     place for two days, then removed. The fruit were checked
                     to be sure the thrips had stayed on, and then brought to
                     the lab to cut off the flavedo using a razor blade. Only
                     the flavedo from the stem 1/3 to 1/2 of the fruit was
                     used. For controls an equal number of comparable sized
                     fruit were caged without thrips, and the peel collected
                     from them in the same manner. Tissues were frozen in
                     liquid nitrogen, then stored at -80C until further
                     processing. Fenton (Close lab) purified RNA by the phenol
                     method described in J. Japanese Soc. Hort. Sci. 1996. 64
                     (4): 809-814, purified poly(A) mRNA using an Oligotex mRNA
                     Kit (Qiagen), produced a primary cDNA library using a
                     lambda ZAP XR cDNA Synthesis Kit (Stratagene), then
                     mass-excised 0.77 million pfu from the primary library to
                     produce a phagemid population. The library was made from
                     equal portions of RNA from each of the two treatments.
                     Phagemids were plated, plasmid DNA purified, cDNA clones
                     archived, and DNA sequences determined bi-directionally

ORIGIN

```
        using an ABI3730 at DNA Landmarks (Landry, Hubert,
        Laforest, Landry, Ligonde). Chromatogram files were
        downloaded by FTP by Close, then processed by Wananmaker.
        (Close lab) using the HarvEST pipeline
        (http://harvest.ucr.edu) to remove vector and cloning
        oligo sequences and various contaminants, and to trim to a
        high quality region. Sequences that retained a phred 17
        region of at least 100 bases were assembled, then chimeras
        were removed following manual inspection of assemblies
        (Close, Roose, Federici, Wananamaker, Lyon, Ye, Jang,
        Collin, Kacar, Ikeda, Quinitio). Sequences that survived
        all removal steps were submitted to GenBank."
```

US-10-019-931-3 (1-359) x CX043920 (1-817)

Alignment Scores:
Pred. No.:

Score:	5.29e-91
	887.50
Percent Similarity:	77.69%
Best Local Similarity:	61.54%
Query Match:	46.76%
DB:	8

Length:	817
Matches:	160
Conservative:	42
Mismatches:	53
Indels:	5
Gaps:	3

```
Qy   30 ThrArgAspValProLysAlaGluAlaGluAspGlyGlyLeuSerCysGluPheGlnGluMet 49
         ||||||||||||:::::|||||||||||||||||:
Db   34 ACAAGAACCCAAATAGCAGATGAAGAAGAAAAGCACTTAAGCAGTGAGTCCAAGGATGTG 93

Qy   50 LeuAspSerLeuProLysGlyGlyArgGlyTyrThrArgTyrLeuTyrLeuPheGlnGlyGly 69
         ||||:||||||||||||||||||||||||||  :::||||||||||||||||||||||||
Db   94 CTTGTTTCAACTTCCTAAAGGAAGGAGGTTGGCGGTTCTTTCTTTATAAGTTCCAGGGG 153

Qy   70 PheTrpCysGlnAlaAlaLysGlyIleGlnAlaIleMetSerPheGlnLysHisPheGlnSer 89
         :::||||||||||||||||||||||||||||||||||:::|||||||||||::::||
Db  154 TATTGGTGCCAGGCGAAAGGAAGGAATCCAATAAGGCTTTCCAAAAGCACTTCAAAAGCA 213

Qy   90 LeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThrTrpLeuTyrLeuAla 109
         :::                 ||||::::||||||||||||||||||||||||||||||
Db  214 AAAGACACTGATATCATTTTAGCCAGCATACCAAAAATCAGGCACCCACTTGGATGAAAGCA 273

Qy  110 LeuThrPheThrIleLeuAsnArgHisArgPheAspProValAlaSer--------Ser 126
         ||||:||| :::||||||||||||||||  :::|||
Db  274 TTGTCTTTGCAATAAATCAATCGCAAAGAATTTC---CCCATAATTAGTGATCATCATGGT 330

Qy  127 ThrAsnHisProLeuPheThrSerAsnProHisAspLeuValProPhePheGluTyrLys 146
         ::::||||||||||||||||||||||||||||||||||||||||
Db  331 CATCATCACCCCCTGCTTACTTCAAAATCCTCATGGATCTGTTCCCTTTTCCTTTGAATACAAG 390

Qy  147 LeuTyrArgAlaAsnGlyAspValProAspLeuSerGlyLeuAla---SerProArgThrPhe 165
         :::|||||||||||||||||||||::::||||||||||||||||||  ||||||||||
Db  391 CTCTATGCAAATAACCAAAATTCCTGAGCTTTCTCAAAATTGCTGATGAGCCTAGAGTTTTT 450

Qy  166 AlaThrHisLeuProPheGlySerLeuLysGluThrIleGluLysProGlyValLysVal 185
         :::|||||:|||
Db  451 GCCACCCATATTCCATTTGGCTTCCTAAAATTGCTTCAATGAACAACATTAAGATCT 510

Qy  186 ValTyrLeuCysArgAsnProPheAspThrPheIleSerSerThrTrpHisTyrThrAsnAsn 205
         ||||:||||:|||||||||||||||:||||||||||||||||
Db  511 GTTTATATTGCAGAAACCCTTTGATACTTTCATCTCTTCATGGCATTTTTCTTAACACAA 570

Qy  206 IleLysSerGluSerValSerProValLeuLeuAspGlnAlaPheAspLeuTyrCysArg 225
         ||:::|||||:::||||||||||||||||||||||||||:||||:::
Db  571 TTGAGGTCTCCAAGGATTCCACTCAGGAGGACCATTCAAAAAGTACTCGTGAT 630

Qy  226 GlyValIleGlyPheGlyProPheTrpGluHisMetLeuGlyTyrThrArgGluSerLeu 245
         ||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db  631 GGTGTCATTGGCGGTTCGGTTCCCATTTGGGAGCATAGTTAGGGTATTGGAATGAGAGCTTG 690

Qy  246 LysArgProGlyLysValPhePheLeuArgTyrGluValSerAspAspIleGluThr 265
         |||||||||:|||||||||||||||:::||||:::|||||||||||||||
Db  691 AAGAGACCAAACAATGTCTTGTTTTGAAGTATACGAATGAAAACAAGACATTGTGTCA 750

Qy  266 AsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrArgGluGluGluArgLys 285
         ||| :::||||||||||||||||||||||||||||||||||||||||||||||
Db  751 AATTTGAAGAAAATTGGCAAGCTTTTTGGGGTTCCCTTTTCCCCTGAGGAGGAGGAACAAG 810
```

Db

RESULT 15
AV827990

LOCUS   AV827990   574 bp   mRNA   linear   EST 01-APR-2002
DEFINITION AV827990 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-22-M04 5',
           mRNA sequence.
ACCESSION AV827990
VERSION   AV827990.1 GI:19870050
KEYWORDS   EST.
SOURCE   Arabidopsis thaliana (thale cress)
  ORGANISM Arabidopsis thaliana
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
           rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 574)
  AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
           Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
           Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
           and Shinozaki,K.
  TITLE   Large scale analysis of Arabidopsis full-length cDNA (2002b)
  JOURNAL Unpublished (2002)
  COMMENT Contact: Motoaki Seki
           Plant Functional Genomics Research Group
           RIKEN Genomic Sciences Center
           3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
           Tel: 81-298-36-4359
           Fax: 81-298-36-9060
           Email: mseki@rtc.riken.go.jp
           An Arabidopsis full-length cDNA library was constructed essentially
           as reported previously (Seki et al., 1998).cDNA clone was ligated
           and XhoI was ligated to modified lambda FLC-1 vector (Carninci et
           al. submitted for publication) digested with BamHI and SalI. This
           clone is in a modified pBluescript vector. Please visit our web
           site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
           details.

FEATURES   Location/Qualifiers
     source   1..574
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="RAFL09-22-M04"
                     /dev_stage="plants at various developmental stages from
                     germination to mature seeds"
                     /lab_host="DH10B"
                     /clone_lib="RAFL9"
                     /note="Site_1: BamHI; Site_2: SalI; subjected to
                     dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
                     hr) treatments"

ORIGIN

US-10-019-931-3 (1-359) x AV827990 (1-574)

Alignment Scores:
Pred. No.:

Score:	3.24e-89
	870.00
Percent Similarity:	98.26%
Best Local Similarity:	98.26%
Query Match:	45.84%
DB:	1

Length:	574
Matches:	169
Conservative:	0
Mismatches:	1
Indels:	2
Gaps:	0

```
Qy    1 MetAlaThrSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHis 20
         ||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGGCTACCTCAAGCATGAAGAGCATTCCAATGGCGGATCCCAAGTTTCTCCATGTGTCAC 63

Qy   21 LysLeuGluLeuLeuLysGluGlyLysThrArgAspValProLysAlaGluGluAspGlu 40
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   63 AAGCTCGAGCTCCTTAAAGAAGGGCAAAACTCGCAGTCCGAAAGCCGAGAAGATGAA 182

Qy   41 GlyLeuSerCysGluPheGlnGluMetLeuAspSerLeuProLysGlyGlyArgGlyTyrArg 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  183 GGGCTAAGCTGCGAGTTCCAAGAGATGTTGGATTCTCTTCCTAAGGAGGAGGAGGATGGAGA 242
```

```
Qy   61  ThrArgTyrLeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGluIleGlnAlaIle   80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  243  ACTCGTTACCTTTACCTTATTCCAAGGGTTTTGGTGCCAAAGCCAAAAGAGATTCAAGCCATC  302

Qy   81  MetSerPheGlnLysHisPheGlnSerLeuGluAsnAspValValLeuAlaThrIlePro  100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  303  ATGTCTTTCCAAAAACATTTCCAATCCCTCGAAAACGACGTCGTTCTCGGCACCATACCT  362

Qy  101  LysSerGlyThrThrTrpLeuLysAlaLeuThrPheThrIleLeuAsnArgHisArgPhe  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  363  AAAATCCGGTACAACCTGGGCTAAAAGCTTTAAACTTTCACCATCCTTAACCGTCACCGGTTT  422

Qy  121  AspProValAlaSerSerThrAsnHisProLeuPheThrSerArgAsnProHisAspLeuVal  140
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  423  GATCCGGTTGCCTCGAGTACCAACCACCCCTCTTTTCACTTCCAACCCTCATGACCTTGTA  482

Qy  141  ProPhePheGluTyrLysLeuTyrAlaAsn-GlyAspValProAspLeuSerGly-LeuA  160
         ||||||||||||||||||||||||||||| ||||||
Db  483  CCTTTCTTCGAGTACAAGCTTTACGGCAACCGGAGATGTTCCGGATCTCTCGNGGTTCTAG  542

Qy  160  1aSerProArgThrPheAlaThrHisLeuPro  170
         |||||| |||||||||||||||||||||||||
Db  543  CCAGTNCAAGAAACGTTCGGCAAACCCACTTACCG  574
```

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 23, 2005, 18:54:37 ; Search time 214 Seconds
                                         (without alignments)
                                         2981.984 Million cell updates/sec

           GenCore version 5.1.6
           Copyright (c) 1993 - 2005 Compugen Ltd.

Title: US-10-019-931-3
Perfect Score: 1898
Sequence: 1 MATSSMKSIPWAIPSFSMCH..........RLSALVDDKLGGSGLITFRLS 359

Scoring table: BLOSUM62
                Xgapop 10.0 , Xgapext 0.5
                Ygapop 10.0 , Ygapext 0.5
                Fgapop 6.0 , Fgapext 7.0
                Delop   6.0 , Delext   7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Command line parameters:
 -MODEL=frame+_p2n.model
 -Q=/cgn2_1/USPTO_spool_p/US10019931/runat_16122005_165219_18469/app_query.fasta_1.519
 -DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LCOPCL=0
 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
 -USER=US10019931 @CGN_1_1_290 @runat_16122005_165219_18469 -NCPU=6 -ICPU=3
 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   Issued_Patents_NA:*
              1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
              2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
              3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
              4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
              5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
              6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
              7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
              8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
              9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

SUMMARIES

Result No.	Score	Pred. No. %	Query Match	Length	DB	ID	Description
1	550.5	29.0	1192	3	US-10-199-329-1	Sequence 15, Appli
2	325.5	17.1	922	3	US-09-609-816-2	Sequence 2, Appli
3	325.5	17.1	922	3	US-10-199-334-2	Sequence 2, Appli
4	325.5	17.1	922	3	US-10-199-329-2	Sequence 2, Appli
5	300	15.8	1004	3	US-09-609-816-3	Sequence 3, Appli
6	300	15.8	1004	3	US-10-199-334-3	Sequence 3, Appli
7	300	15.8	1004	3	US-10-199-329-3	Sequence 3, Appli
8	296.5	15.6	942	3	US-09-609-816-1	Sequence 1, Appli
9	296.5	15.6	942	3	US-10-199-334-1	Sequence 1, Appli

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

No.	Score	Length	DB	ID	Description
10	296.5	942	3	US-10-199-329-1	Sequence 2, Appli
11	294	1396	3	US-09-328-17AA-2	Sequence 4457, Appli
12	286.5	1045	3	US-09-949-016-4457	Sequence 655, App
13	286.5	1063	3	US-08-225-562-1	Sequence 1, Appli
14	286.5	1063	2	US-08-437-795-1	Sequence 45, Appli
15	278.5	993	3	US-09-717-321A-45	Sequence 1, Appli
16	275	912	3	US-09-795-926-1	Sequence 5, Appli
17	275	912	3	US-10-364-774-1	Sequence 5, Appli
18	266.5	798	3	US-09-795-926-5	Sequence 5, Appli
19	266.5	798	3	US-10-364-774-5	Sequence 147, App
20	263	2431	3	US-09-949-016-147	Sequence 5340, Ap
21	244.5	1447	3	US-09-949-016-5340	Sequence 655, App
22	244.5	1477	3	US-09-949-016-655	Sequence 22, Appli
23	241	1130	3	US-09-795-926-22	Sequence 19, Appli
24	239.5	2153	3	US-09-795-926-19	Sequence 19, Appli
25	239.5	2153	3	US-10-364-774-19	Sequence 15, Appli
26	227.5	645	3	US-09-795-926-15	Sequence 15, Appli
27	227.5	645	3	US-10-364-774-15	Sequence 2593, App
28	209	1196	3	US-09-949-016-2593	Sequence 1, Appli
29	197	1564	2	US-08-852-481-1	Sequence 11, Appli
30	144.5	561	3	US-09-795-926-11	Sequence 11, Appli
31	144.5	561	3	US-10-364-774-11	Sequence 7, Appli
32	136	447	3	US-09-795-926-7	Sequence 7, Appli
33	136	447	3	US-10-364-774-7	Sequence 340, App
34	113.5	289	3	US-09-490-609B-340	Sequence 340, Appli
35	109.5	4411529	3	US-09-103-840A-1	Sequence 16, Appli
36	108.5	304	3	US-09-172-711-16	Sequence 24, Appli
37	108.5	978	3	US-10-126-229-24	Sequence 24, Appli
38	108.5	978	3	US-10-286-606-24	Sequence 2, Appli
39	106.5	4403765	3	US-09-103-840A-2	Sequence 2420, Ap
40	103.5	924	3	US-09-328-352-2420	Sequence 8, Appli
41	103.5	1485	2	US-08-376-843-8	Sequence 8, Appli
42	103.5	1485	2	US-08-468-036-8	Sequence 1, Appli
43	103.5	3098	2	US-08-447-500-1	Sequence 1, Appli
44	103.5	3098	2	US-08-454-097-1	Sequence 1, Appli
45	103.5	3098	2	US-08-447-408-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-854-122-15
; Patent No. 6841718
; Sequence 15, Application US/09854122
; GENERAL INFORMATION:
; APPLICANT: ALBERTE, RANDALL S.
; APPLICANT: SMITH, ROBERT
; TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
; FILE REFERENCE: PHA-007.01
; CURRENT APPLICATION NUMBER: US/09/854,122
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/202,529
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 1192
; TYPE: DNA
; ORGANISM: Zostera marina
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1..33, 37..1041, 1045..1059, 1063..1077,
; LOCATION: 1081..1086, 1090..1119, 1123..1191)
US-09-854-122-15

Alignment Scores:
Pred. No.:
Score: 9.38e-57
Pred. No.:   550.50
Percent Similarity: 57.91%
Best Local Similarity: 36.71%
Query Match: 29.00%
DB:

Length: 1192
Matches: 116
Conservative: 67
Mismatches: 122
Indels: 11
Gaps: 8

US-10-019-931-3 (1-359) x US-09-854-122-15 (1-1192)

```
Qy   30 ThrArgAspValProLysAlaGlyLeuAspGlyLeuSerCysGluPheGlnGlnMet  49
              :::       |||||||
Db   85 TCCAAGAATGAGCAAGAAGAAGGAAGAAGAATCCAAAATGACAGAGAGATATAGAGAGAATT 144

Qy   50 LeuAspSerLeuProLysGluArgGlyTrpArgThrArgTyrLeuPheGlnGly  69
        |||
Db  145 GTTTCTTCACTTCCCTGAATCATTATTGGGGGGATACC---ATGAGGTTGTACAAGGGA 201

Qy   70 PheTrpCysGlnAlaLysGluIleGlnAlaIleMetSerPheGlnLysHisPheGlnSer  89
              |||:::::||        :::
Db  202 CGAGAGACGGACATTATCCTTACGACTCTTCCAAAGGCTGGAACGACACATGGACGAAGGCA 261

Qy   90 LeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThrTrpLeuLysGluAla 109
                 |||:::|||  ::::::||
Db  262 CTGAGACGTTTGCCATCCTTAACACAGATGTTAACCACACCA----TCATCACCGACACAT 321

Qy  110 LeuThrHisPheThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThrAsnHis 129
             ||||::::::|||    :::
Db  322 CTGACGTTTGCCATCCTTAACACAGATGTTAACCACACCA------TCATCACCGACACAT 375

Qy  130 ProLeuPheThrSerAsnProHisAspLeuValProPhePheGluThrTyrLysLeuTyrArgAla 149
             |||:::|||   |||
Db  376 CCACTTTGTTCTTCAACCCTCATGATCTTGTTCCTTTCTTTGAAACTTACAAAATTATAT 432

Qy  150 AsnGlyAsp------ValProAlaProLeuSerGlyLeu 166
             :::  ||||||
Db  433 GGTAGAGAAAATACGATGCCAGACCTCGAATATGTTGAATGAATCGCCGAGGTTGTTTGCC 492

Qy  167 ThrHisIleLeuProPheGlySerLeuValGluThrIleGluLysGlyValLysValValVal 186
             |||   :::
Db  493 GGACAATCCCCATACTCTTTGGTTGCCGCGTCGTGTTTTGAAATCGGGAACAAAAATCATC 552

Qy  187 TyrLeuCysArgAsnProPheAspThrPheIleSerSerTrpHisTyrThrAsnAsnIle 206
             |||:::|||
Db  553 AATATAAGCCGGCAAACCCGTAAGAGTACATTTGTGTCTTTTTGGAAAATTCGCAATCTGATT 612

Qy  207 LysSerGluSerValSerProValLeuLeuAspLeuIleLysGlyLeuProTyrCysArgGly 226
             |||:::       :::
Db  613 AACCCCGACAAGTTA------TTGGACCTGGAAAGACGGTTGAATAATATCTTCGGCATCGGGA 666

Qy  227 ValIleGlyPheGlyProProPheTrpGluHisMetLeuGlyTyrTrpArgGluThrLys 246
             |||:::
Db  667 ATCTCCTTTGTGGACCGGAATGGAATTTCCAAGCGGAGTTC---ACCAATGCGCGTCT 723

Qy  247 ArgProGlyLysValPhePheLeuArgTyrGluAspAspIleGluThrArgAsn 266
        |||             :::
Db  724 ACTAAATTCAAACTTGCTATTGTCATTGTGAGTTGTAGGAAACCCAGTTGAAAAT 783

Qy  267 LeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGluGluLeuArgGlyGly 286
             :::|||||||
Db  784 GTGAAGAAGCTAGCTGAGTTCATCGGGGATTCAGAGAGTGAGGAGAAACAGGG 843

Qy  287 ValValLysAlaIleAlaGluLeuCysSerPheGluAsnLeuLysLysLeuGluValAsn 306
             |||:::|||
Db  844 ATTGTTGATGAGATAGTTTAAACTTTGTAGCTTGGAGAATCAACAGTGAAC 903

Qy  307 LysSerArgAsnLysSerIleLysLysAsn---PheGluLeuProPheLeuArgLysGlyGlu 325
             :::  :::|||
Db  904 AAAAACGGATCAAGCTACAATTCGAAAATCGACAACAAGCCATTTCTTCAGGAAAGGTGAG 963

Qy  326 ValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArgLeu 341
             |||:::
Db  964 GTGAGAGATTGGGCAACTATCTACGTCGGGAAATGATTAAGAAAACTG 1011
```

Alignment Scores:

```
Pred. No.:                        1.4e-29
Score:                            325.50
Percent Similarity:   46.94%      Length:          922
Best Local Similarity: 29.25%     Matches:         86
Query Match:          17.15%      Conservative:    52
DB:                   3           Mismatches:      101
                                  Indels:          55
                                  Gaps:            8
```

US-10-019-931-3 (1-359) x US-09-609-816-2 (1-922)

```
Qy   86 HisPheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThr 105
             :::|||||||:::
Db  128 AATTTCCAAGCCAAGCCTGATGATCTTAATTCTCGGCAACTTACCCAAAGTCAGGTACAACA 187

Qy  106 TrpLeuLysGluAlaLeuThrPheThrIleLeuAsn--------------- 116
             |||||||:::
Db  188 TGGATGCATGAAAATTTTAGACATGATTCTAAATGATGGTGGAGAAATGCAAAAGA 247

Qy  117 ------ArgHisArgPheAspProValAlaSerSerThrAsnHisProLeu 131
                  :::|||
Db  248 GCCCAGACTCTAGATAGACACGCTTC----------------CTT 277

Qy  132 PheThrSerAsnProHisAspLeuValProPhePheGluThrTyrArgLeuTyrArgAlaAsnGly 151
             ||||:::
Db  278 GAACTGAAATTTCCCCATAAAGAAAAAACCAGATTTGGAGTTCGGTTCTT------- 325

Qy  152 AspValProLeuSerGlyLeuValAlaSerProArgThrPheAlaThrHisLeuProPhe 171
             :::::::|||
Db  326 ------GAAATGTCCTCCACCAACATCTATAAAGAAAAAACCAGATTTGGAGTTCGGTTCTT 367

Qy  172 GlySerLeuValGlyThrIleGluLysGlyValLysValAlaValTyrLeuCysArgAsn 191
             |||:::|||
Db  368 CATCTGATTCCACCATCTATCTGGAAAGAAAAACTGCAAGAATTGTCTATGTGGCCAGAAAT 427

Qy  192 ProPheAspThrPheIleSerSerTrpHisTyrThrAsnAsnIleLysSerGluSerVal 211
             |||  :::
Db  428 CCCAAGGATTGCCTGGTGTTCCTACTTACCACACTTT---CACAGAGTAGGCTTCCTTTATGCCT 484

Qy  212 SerProValLeuLeuAspLeuIleLysGlyLeuProTyrCysArgGlyValIleGlyPheGly 231
             :::   |||
Db  485 GATCCTCAGAACTTAGAGAAATTTTTATGAGAAATTCATGTCCGGAAAAAGTTGTTGGCGGG 544

Qy  232 ProProPheTrpGluHisMetLeuGlyTyrTrpArgGluThrArgAsnLeuLysArgLeuAla 251
             |||   :::
Db  545 TCCTTGGTTTGACCAATGTGAAAGAGGGTGG------GCTGCAAAAGACAATGCACCGGATC 598

Qy  252 PhePheLeuArgTyrGluAspAspIleGluThrArgAsnLeuLysArgLeuAla 271
             :::|||::::::|||
Db  599 CTCTACCTCTTCTACGAGGATATTAAAAAGACCCAAGCGGGAAATTGAGAAGATACTG 658
```

RESULT 3

US-10-199-334-2

; Sequence 2, Application US/1019934
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 922
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-199-334-2

Alignment Scores:

Pred. No.:	
Score:	325.50
Percent Similarity:	46.94%
Best Local Similarity:	29.25%
Query Match:	17.15%

US-10-019-931-3 (1-359) x US-10-199-334-2 (1-922)

Length:	922	
Matches:	86	
Conservative:	52	
Mismatches:	101	
Indels:	55	
Gaps:	8	

RESULT 4

US-10-199-329-2

; Sequence 2, Application US/10199329
; Patent No. 6953681
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 922
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-199-329-2

Patent No. 6436684

GENERAL INFORMATION:
APPLICANT: Woodage, Trevor
APPLICANT: Wei, Minh Hui
APPLICANT: Kodira, Chinnappa
APPLICANT: Beasley, Ellen
APPLICANT: DiFrancesco, Valentina
TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
FILE REFERENCE: CL000669PCT
CURRENT APPLICATION NUMBER: US/09/609,816
CURRENT FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/192,408
PRIOR FILING DATE: 2000-03-27
PRIOR APPLICATION NUMBER: 60/212,725
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/609,816
PRIOR FILING DATE: 2000-07-03
NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 3
  LENGTH: 1004
  TYPE: DNA
  ORGANISM: HUMAN
US-09-609-816-3

Alignment Scores:
Pred. No.:
Score: 300.00
Percent Similarity: 42.82%
Percent Similarity:
Best Local Similarity: 25.75%
Query Match: 15.81%
DB:

Length: 1004
Matches: 95
Conservative: 63
Mismatches: 121
Indels: 90
Gaps: 14

US-10-019-931-3 (1-359) x US-09-609-816-3 (1-1004)

```
Qy   16 PheSerMetCysHisLysLeuGluLeuLeuLeuLeuGluGlyLysGlyLysThrArgAspValProLys 35
            ::|||:::
Db   38 TTCGCCCTTATGCGGGAAGATTGAG------------------AAAAACGTCGCCACG 76

Qy   36 AlaGluGluAspGluGlyGlyLeuSerCysGlyPheGlnGluMet------LeuAspSerLeuPro 54
            |||:::
Db   77 ATGGAAAAAAGCCAGAACTG------TTTAACATCATGGAAGTAGATGGAGTCCCT 127

Qy   55 LysGluArgGlyGlyTrpArgThrArgTyrLeuTyrLeuPheGlnGlyPheTrp------ 71
            |||:::
Db  128 ACG------------------------TTGATATTATCAAAAGAATAGTGGGGAAAAGTC 163

Qy   72 CysGlnAlaLysGlyGluIleGlnAlaIleMetSerPheGlnLysHisLeuSerLeuGlu 91
            :::|||||||:::
Db  164 TGT------------------AATTTCCAAGCCAAGCCT 184

Qy   92 AsnAspValValLeuAlaThrIleProLysSerGlyThrThrTrpLeuLysAlaLeuThr 111
            ::|||||||
Db  185 GATGATCTTATTCTGGCAACTTACCCAAGTCAGGTACAACATGGATGAAATTTTA 244

Qy  112 PheThrIleLeuAsn------------------------ 117
                                                 Arg
Db  245 GACATGATTCTAAATGATGGTGATGTGATGGAGAAATGCAAAGAGCCCAGACTCTAGATAGA 304

Qy  118 HisArgPheAspProValAlaSerSerThrAsnHisProHis 137
            :::|||::
Db  305 CACGCTTTC------------------------CTTGAACTGAAATTTCCCAT 334

Qy  138 AspLeuValProPhePheGluTyrLysLeuTyrAlaAsnGlyAspValProAspLeuSer 157
            |||
Db  335 AAAGAAAAAACCAGATTTGAGTTCGTTCTT------------------ 364

Qy  158 GlyLeuAlaSerProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThr 177
            ::::|||||||:::
Db  365 GAAATGTCCTCACCACACTGATAAAAAACACATCTCCCTTCACATCTGATTCCACCATCT 424
```

Alignment Scores:
Pred. No.:   1.4e-29
Score: 325.50
Percent Similarity: 46.94%
Percent Similarity:
Best Local Similarity: 29.25%
Query Match: 17.15%
DB:

Length: 922
Matches: 86
Conservative: 52
Mismatches: 101
Indels: 55
Gaps: 8

US-10-019-931-3 (1-359) x US-10-199-329-2 (1-922)

```
Qy   86 HisPheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThr 105
            ::|||||||
Db  128 AATTTCCAAGCCAAGCCTGATCGATCTTATTCTGGCAACTTACCCAAGTCAGGTACAACA 187

Qy  106 TrpLeuLysAlaLeuThrPheThrIleLeuAsn------------------------ 116
            |||:::
Db  188 TGGATGCATGAAATTTTAGACATGATTCTAAATGATGGTGATGTGATGGAGAAATGCAAAAGA 247

Qy  117 ------------ArgHisArgPheAspProValAlaSerSerThrAsnHisProLeu 131
                       CTT
Db  248 GCCCAGACTCTAGATAGACACGCTTTC------------------------ 277

Qy  132 PheThrSerAsnProHisAspLeuValProPhePheGluTyrLysLeuTyrAlaAsnGly 151
            ||||||
Db  278 GAACTGAAATTTCCCATAAAGAAAAACCAGATTTGAGTTCGTTCTT 325

Qy  152 AspValProAspLeuSerGlyLeuAlaSerProArgThrPheAlaThrHisLeuProPhe 171
            |||||||:
Db  326 ------------GAAATGTCCTCACCACACTGATAAAAAACACATCTCCCTTCA 367

Qy  172 GlySerLeuLysGluThrIleGluLysProGlyValLysValValThrTyrLeuCysArgAsn 191
            |||::|||:
Db  368 CATCTGATTCCACCATCTATCTGGAAAGATGCAAGATTGTCTATGTGGGCCAGAAAT 427

Qy  192 ProPheAspThrPheIleSerSerTrpThrAsnLeuIleLysSerGluSerVal 211
            |||:|||
Db  428 CCCAAGGATTGCCTGGTGTCCTACTACCACTTT 484

Qy  212 SerProValLeuLeuAspGlnAlaPheAspLeuTyrCysGlyArgGlyValIleGlyPheGly 231
            |||||||:
Db  485 GATCCTCAGAACTCTAGAGGAATTTATCAGTCCCGGAAAAGTTGTTGGCGGG 544

Qy  232 ProPheTrpGluHisMetLeuGlyTyrArgTrpArgSerArgProGluLysVal 251
            :::
Db  545 TCCTGGTTTGACCATGTGAAAGGATGGTGG 598

Qy  252 PhePheLeuArgTyrGluAspLeuLysSerAspArgIleGluThrArgAsnLeuLysSerArgLeuAla 271
            :::|||
Db  599 CTCTACCTCTTCTTACGAGGATATTAAAAAAAGACCCAAAGCGGGAAAATTGAGAAGATACTG 658

Qy  272 ThrPheLeuGluLeuProPheThrGluGlyLeuGlyLeuArgLysGlyValValLysAlaIle 291
            |||||||
Db  659 AAGTTCCTGGAAAAAGACATATCAGAGGAA 703

Qy  292 AlaGluLeuCysSerPheLeuAsnLeuLysLys------------ 303
                                              Leu
Db  704 ATCTATCACACCTCCTTTGATGTAATGAGCAACAATGACCAACTATACCACTTTG 763

Qy  304 GluValAlanLysSerAsnLysSerIleLysAsnPheGluLeuPheArgLys 323
            :::
Db  764 CCCACCAGCATTATGGACACCATCCATCTCCCCTTTT 808

Qy  324 GlyGluValSerAspTrpValAsnTyrLeuSerProSerGlnValGlyArgLeuSerAla 343
            |||
Db  809 GGGAATGCCTGGAGACTGGAAGACTATTTTACGTGTGGGCCCAAAATGAAGAATTTGACAAG 868

Qy  344 LeuValAspAspAspLysGlyGlySerGlyLeuGlyLeuThrPheArg 357
            |||::|||
Db  869 GACTACCAGAAGAGAGAATGGCAGGGAGGAAGCACCCTAACCTTCCGC 910
```

RESULT 5
US-09-816-3
; Sequence 3, Application US/0960816

RESULT 6
US-10-199-334-3
; Sequence 3, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1004
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-199-334-3

Alignment Scores:
Pred. No.:
Score: 2.0e-26
Score: 300.00
Percent Similarity: 42.82%
Best Local Similarity: 25.75%

Length: 1004
Matches: 95
Conservative: 63
Mismatches: 121

US-10-019-931-3 (1-359) x US-10-199-334-3 (1-1004)

Query Match:
DB: 3
                15.81%
Indels: 90
Gaps: 14

[Sequence alignment blocks — Qy (query) and Db (database) paired rows with alignment bars]

```
Qy    330  ValAsnTyrLeuSerProSerGlnValGluArgLeuSerAlaLeuValAspAspLysLeu 349
Db    866  AAGAACCACTTTACTGTGGCTATGAAGTGAACTTTGATAAGCATTATGAAAAGAAGATG 925

Qy    350  GlyGlySerGlyLeuThrPheArgLeu 358
Db    926  GCAGGGTCCACACTGAACTTCTGCCTG 952
```

RESULT 7
US-10-199-329-3
```
; Sequence 3, Application US/1019329
; Patent No. 6953681
; GENERAL INFORMATION:
;  APPLICANT: WOODAGE, Trevor et al.
;  TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
;  TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
;  TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL000669DIV-2
;  CURRENT APPLICATION NUMBER: US/10/199,329
;  CURRENT FILING DATE: 2002-07-22
;  PRIOR APPLICATION NUMBER: 09/609,816
;  PRIOR FILING DATE: 2001-03-27
;  PRIOR APPLICATION NUMBER: 60/192,408
;  PRIOR FILING DATE: 2000-03-27
;  PRIOR APPLICATION NUMBER: 60/212,725
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: 09/609,816
;  PRIOR FILING DATE: 2000-07-03
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 3
;  LENGTH: 1004
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-10-199-329-3
```

Alignment Scores:
```
Pred. No.:
Score:                 2.01e-26
                       300.00
Percent Similarity:    42.82%
Best Local Similarity: 25.75%
Query Match:           15.81%
DB:                    3

Length:         1004
Matches:        95
Conservative:   63
Mismatches:     121
Indels:         90
Gaps:           14
```

US-10-019-931-3 (1-359) x US-10-199-329-3 (1-1004)

```
Qy    16  PheSerMetCysHisLysLysGluLeuLeuLysGluGlyLysThrArgAspValProLys 35
Db    38  TTCGCCCTTATGGCGAAGAAGATGAG-----------------AAAAACGCTCCACG 76

Qy    36  AlaGluGluAspGlyGlyLeuSerCysGluPheGlnGluMet---LeuAspSerLeuPro 54
Db    77  ATGGAAAAAAGCCAGAACTG----TTTAACATCATGGAAGTAGATGGAGTCCCT 127

Qy    55  LysGluArgGlyTrpArgThrArgTyrTyrLeuPheGlnGlnTrp------- 71
Db    128 ACG--------TTGATATTATCATCAAAAAGGAATGGTGGGAAAAGTC 163

Qy    72  CysGlnAlaLysGluGluIleGlnIleAlaIleMetSerPheGlnLysHisPheGlnSerLeuGlu 91
Db    164 TGT--------AATTTCCAAGCCAAGCCT 184

Qy    92  AsnAspValValLeuAlaThrIleProLysSerGlyTyrThrThrThrLeuValSerAlaLeuThr 111
Db    185 GATGATCTAATTCCTGCAACTTACCCAAAGTCAGGTACACATGGAAGTGCATTA 244

Qy    112 PheThrIleLeuAsn------ 117
Db    245 GACATGATTCTAAAATGATGGTGGTGAGAAATGCAAAAGAGCCCAGACTCTAGATAGA 304

Qy    118 HisArgPheAspProValAlaSerSerThrAsnHisProLeuPheThrSerAsnProHis 137
Db    305 CACGCTTTC-------------------CTTGAAACTGAAATTTCCCAT 334
```

```
Qy    138 AspLeuValProPhePheGluTyrLysLeuTyrTyrAlaAsnGlyAspValProAspLeuSer 157
Db    335 AAAGAAAAACCAGATTGGAGTTCGTTCTT----------------------- 364

Qy    158 GlyLeuAlaSerProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThr 177
Db    365 GAAATGTCCTCACCACCACTTT--CACAGGATGGCTTCCTTATGCCTGTAACTTAGA 424

Qy    178 IleGluGlyProGlyValLysValValTyrLeuCysArgAsnProPheAspThrPheIle 197
Db    425 ATCTGGAAGAAAACTGCAAGATTGTCTTA-TGTGGCCAGAAAATCCAAGGATTGCCTGGTG 484

Qy    198 SerSerTrpHisTyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAsp 217
Db    485 TCCTACTACCACTTT--CACAGGATGGCTTCCTTATGCCTGTAACTTCCAGAACTTAGA 541

Qy    218 GlnIleAlaPheAspLeuTyrCysSerArgGlyValIleGlyProPheProTrpGlnHisMet 237
Db    542 GAATTTTATGAGAAATTCATGTCCGGAAAAGTTGTTGGCGAGGTCCTCGGTTTGACCATGTG 601

Qy    238 LeuGlyTyrTrpArgGluSerLeuLeuLysArgProGluLysValPhePheLeuArgTyrGlu 257
Db    602 AAAGGATGGTGG------GCTGCAAAAGGACACGCCACCGGATCCTCTACCTCTTCTTACGGAG 655

Qy    258 AspLeuLysAspAspIleGluThrGlnAsnLeuLeuLysArgLeuAlaThrPheLeuGluLeuPro 277
Db    656 GATATTAAAAAAATCCAAAACATGAAGATCCACAAGGTGTTGGAAATTCTTGGAGAAAACT 715

Qy    278 PheThrGluGluGluGluGluArgLysGlyValValAlaIleAlaGluLeuCysSerPhe 297
Db    716 TTGTCAGGTGAT-----------GTTATAAACAAGATTGTCCACCATACCTCATTT 760

Qy    298 GluAsnLeuLys------------ 309
Db    761 GATGTAATGAAGGATAAATCCATGGCCAACCATACTGCGGTACCTGCTTCCACATATTCAAT 820

Qy    310 LysGluSerIleLeuLysAsnPheLeuGluAsnArgPheArgLysGlyGlyValSerTrp 329
Db    821 CACTCCATCTCAAAATTT------ATGAGGAAAGGGAGGATGCCTGGAGACTGG 865

Qy    330 ValAsnTyrLeuSerProSerGlnValGluArgLeuSerAlaLeuValAspAspLysLeu 349
Db    866 AAGAACCACTTTACTGTGGCTATGAAGTGAACTTTGATAAGCATTATGAAAAGAAGATG 925

Qy    350 GlyGlySerGlyLeuThrPheArgLeu 358
Db    926 GCAGGGTCCACACTGAACTTCTGCCTG 952
```

RESULT 8
US-09-609-816-1
```
; Sequence 1, Application US/09609816
; Patent No. 6436684
; GENERAL INFORMATION:
;  APPLICANT: Woodage, Trevor
;  APPLICANT: Wei, Minh Hui
;  APPLICANT: Kodira, Chinnappa
;  APPLICANT: Beasley, Ellen
;  APPLICANT: DiFrancesco, Valentina
;  TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
;  TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
;  TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
;  FILE REFERENCE: CL000669PCT
;  CURRENT APPLICATION NUMBER: US/09/609,816
;  CURRENT FILING DATE: 2000-07-03
;  PRIOR APPLICATION NUMBER: 60/192,408
;  PRIOR FILING DATE: 2000-03-27
;  PRIOR APPLICATION NUMBER: 60/212,725
;  PRIOR FILING DATE: 2000-06-20
;  PRIOR APPLICATION NUMBER: 09/609,816
;  PRIOR FILING DATE: 2000-07-03
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
```

Left column

```
; SEQ ID NO 1
; LENGTH: 942
; TYPE: DNA
; ORGANISM: HUMAN
US-09-609-816-1

Alignment Scores:
Pred. No.:
Score:                   4.83e-26
                         296.50
Percent Similarity:      45.76%
Best Local Similarity:   27.46%
Query Match:             15.62%
DB:

                         Length:        942
                         Matches:       81
                         Conservative:  54
                         Mismatches:    105
                         Indels:        55
                         Gaps:          8

US-10-019-931-3 (1-359) x US-09-609-816-1 (1-942)
```

```
Qy   86  HisPheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThr 105
         :::|||||||:::
Db  128  AATTTCCAAGCCAAGCCTGATGAATCTTATTCTGGCAACTTACCCAAAGTCAGGTACAACA 187

Qy  106  TrpLeuLysAlaLeuThrPheThrIleLeuAsn------------------------ 116
         |||:|||
Db  188  TGGATGCATGAAATTTTAGACATGATTCTAAATGATGGTGGAGAAATGCAAAAGA 247

Qy  117  ------------ArgHisArgPheAspProValAlaSerSerThrAsnHisProLeu 131
Db  248  GCCCAGACTCTAGATAGACACGGCTTTC------------------------CTT 277

Qy  132  PheThrSerAsnProHisAspLeuValProPhePheGluTyrLysLeuTyrAlaAsnGly 151
         |||:|||
Db  278  GAACTGAAATTTCCCCATAAAGAAAAAAACCAGATTTGGAGTTCGTTCTT-------- 325

Qy  152  AspValProAspLeuSerGlyLeuAlaSerProArgThrPheAlaThrHisLeuProPhe 171
         |||||||||||:::
Db  326  ------GAAATGTCCTCACCACAACTGATAAAAACATCTCCCTTCA 367

Qy  172  GlySerLeuLysGluThrIleGluLysProGlyValLysValValTyrLeuCysArgAsn 191
         |||||||||||
Db  368  CATCTGATTCCACCATCTATCTGGAAAGAAAACTGCAAGATTGTCTATGTGGCCAGAAAT 427

Qy  192  ProPheAspThrPheIleSerSerTrpHisTyrThrArgAsnIleLysSerGluSerVal 211
         |||
Db  428  CCCAAGGATTGCCTCGGTTGTCCTACTACCACACTTT------CACAGGATGGCTTC 484

Qy  212  SerProValLeuLeuAspGlnAlaPheHisAspLeuTyrCysArgGlyValIleGlyPheGly 231
         |||||||||:::
Db  485  GATCCTCAGAACTTAGAGGAATTTATGAGAAATTCAGTCCGGAAAAGTGTTGTGGCGGG 544

Qy  232  ProPheTrpGluHisMetLeuGlyTyrThrArgGlnSerLeuLysArgProGlyLysVal 251
         :::|||||:::|||
Db  545  TCCTGGTTTGCACCATGTGAAAGGATGGTGG------GCTGCAAAAGACATGCACCGATC 598

Qy  252  PhePheLeuArgTyrGluAspLeuLysAspAspIleGluThrAsnLeuLysArgLeuAla 271
         :::|||||||||:::
Db  599  CTCTACCTCTTCTACGAGGATATTAAAAAATCCAAAACATGAGATCACAAAGGTGTTG 658

Qy  272  ThrPheLeuGluLeuProPheThrGluGluGluLeuGluArgLysGlyValValLysAlaIle 291
         :::|||||||
Db  659  GAAATTCTTGGAGAAAAACTTGGTCAGGTGAT------GTTTATAAACAAGATT 703

Qy  292  AlaGluLeuCysSerPheGluAsnLeuLys------------LysLeu 303
Db  704  GTCCACCATACCTCATTTGATGTAAGAAGGATAATCCCATGGCCAACCATACTGCGGTA 763

Qy  304  GluValAsnLysSerAsnLysSerIleLysLysAsnArgPheLeuPheArgLys 323
         |||
Db  764  CCTGCTCACATATTCAATCACTCCAATCCAAAATTT------ATGAGGAAA 808

Qy  324  GlyGluValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArgGlnSerAla 343
         :::
Db  809  GGGAATGCCTGGAAGAAGACCACTTTACTGGTGGCTTTGAATGAGAACTTTGATAAG 868

Qy  344  LeuValAspAspLysLeuGlyGlySerGlyLeuThrPheArgLeu 358
```

Right column

```
Db  869 CATTATGAAAAGAAGATGGCAGGGGTCCAACTGAACTTCTGCCTG 913
        :::        ||||:::     |||||||      |||    |||    |||
```

```
US-10-199-334-1
RESULT 9
US-10-199-334-1
; Sequence 1, Application US/10199334
; Patent No. 6905855
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-3
; CURRENT APPLICATION NUMBER: US/10/199,334
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-199-334-1

Alignment Scores:
Pred. No.:
Score:                   4.83e-26
                         296.50
Percent Similarity:      45.76%
Best Local Similarity:   27.46%
Query Match:             15.62%
DB:

                         Length:        942
                         Matches:       81
                         Conservative:  54
                         Mismatches:    105
                         Indels:        55
                         Gaps:          8

US-10-019-931-3 (1-359) x US-10-199-334-1 (1-942)
```

```
Qy   86  HisPheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThr 105
         :::|||||||:::
Db  128  AATTTCCAAGCCAAGCCTGATGAATCTTATTCTGGCAACTTACCCAAAGTCAGGTACAACA 187

Qy  106  TrpLeuLysAlaLeuThrPheThrIleLeuAsn------------------------ 116
         |||:|||
Db  188  TGGATGCATGAAATTTTAGACATGATTCTAAATGATGGTGGAGAAATGCAAAAGA 247

Qy  117  ------------ArgHisArgPheAspProValAlaSerSerThrAsnHisProLeu 131
Db  248  GCCCAGACTCTAGATAGACACGGCTTTC------------------------CTT 277

Qy  132  PheThrSerAsnProHisAspLeuValProPhePheGluTyrLysLeuTyrAlaAsnGly 151
         |||:|||
Db  278  GAACTGAAATTTCCCCATAAAGAAAAAAACCAGATTTGGAGTTCGTTCTT-------- 325

Qy  152  AspValProAspLeuSerGlyLeuAlaSerProArgThrPheAlaThrHisLeuProPhe 171
         |||||||||||:::
Db  326  ------GAAATGTCCTCACCACAACTGATAAAAACATCTCCCTTCA 367

Qy  172  GlySerLeuLysGluThrIleGluLysProGlyValLysValValTyrLeuCysArgAsn 191
         |||||||||||
Db  368  CATCTGATTCCACCATCTATCTGGAAAGAAAACTGCAAGATTGTCTATGTGGCCAGAAAT 427

Qy  192  ProPheAspThrPheIleSerSerTrpHisTyrThrArgAsnIleLysSerGluSerVal 211
         |||
Db  428  CCCAAGGATTGCCTCGGTTGTCCTACTACCACACTTT------CACAGGATGGCTTC 484

Qy  212  SerProValLeuLeuAspGlnAlaPheHisAspLeuTyrCysArgGlyValIleGlyPheGly 231
         |||||||||:::
Db  485  GATCCTCAGAACTTAGAGGAATTTATGAGAAATTCAGTCCGGAAAAGTGTTGTGGCGGG 544
```

RESULT 10
US-10-199-329-1
; Sequence 1, Application US/1019329
; Patent No. 6953681
; GENERAL INFORMATION:
; APPLICANT: WOODAGE, Trevor et al.
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000669DIV-2
; CURRENT APPLICATION NUMBER: US/10/199,329
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/192,408
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/212,725
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/609,816
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 1
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-199-329-1

SOFTWARE: FastSEQ for Windows Version 4.0

US-10-019-931-3 (1-359) x US-10-199-329-1 (1-942)

Alignment Scores:
Pred. No.:
Score: 296.50
Percent Similarity: 45.76%
Best Local Similarity: 27.46%
Query Match: 15.62%
DB:

Length: 942
Matches: 81
Conservative: 54
Mismatches: 105
Indels: 55
Gaps: 8

Score: 4.83e-26

Qy   232 ProPheTrpGluHisMetLeuGlyTyrTrpArgGluSerLeuLysArgProGluLysVal 251
Db   545 TCCTGGTTTGACCATGTGAAAGGATGGTGG------GCTGCAAAAGACATGCACCGGATC 598

Qy   252 PhePheLeuArgTyrGluAspLeuLysAspAspIleGluThrAsnLeuLysArgLeuAla 271
Db   599 CTCTACCTCTTCTACGAGGATATTAAAAAAAATCCAAAACATGAGATCCACAAGGTGTTG 658

Qy   272 ThrPheLeuGluLeuProPheThrGluGluGluArgLysGlyValValLysAlaIle 291
Db   659 GAATTCTTGGGAGAAAAACTTGGTCAGGTGAT------GTTATAAACAAGATT 703

Qy   292 AlaGluLeuCysSerPheGluAsnLeuLys 303
Db   704 GTCCACCATACCTCATTTGATGTAATGAAGGATAATCCCATGCCAACCATACTGCGGTA 763   LysLeu

Qy   304 GluValAsnLysSerAsnLysSerIleLysSerAsnArgPheLeuPheArgLys 323
Db   764 CCTGCTCACATATTCAATCACTCCATCTCAAAATTT------ATGAGGAAA 808

Qy   324 GlyGluValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArgLeuSerAla 343
Db   809 GGGATGGCTGGAGACTGGAAGAACCACTTACGTGGCTTTGAATGAGAAACTTTGATTAAG 868

Qy   344 LeuValAspAspLysLeuGlyLySerGlyLeuThrPheArgLeu 358
Db   869 CATTATGAAAGAAGATGGCAGGGTCCACACTGAACTTCTGCCTG 913

RESULT 11
US-09-328-174A-2
; Sequence 2, Application US/09328174A
; Patent No. 6448003
; GENERAL INFORMATION:
; APPLICANT: Guida, Marco
; APPLICANT: Kurth, Janice
; TITLE OF INVENTION: Genotyping Human Phenol Sulfotransferase
; FILE REFERENCE: 4389-6 (formerly SEQ-16P)
; CURRENT APPLICATION NUMBER: US/09/328,174A
; CURRENT FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 09/328,174
; PRIOR FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1396
; TYPE: DNA
; ORGANISM: H. sapiens

Qy   117 ------------------ArgHisArgPheAspProValAlaSerThrAsnHisProLeu 131
Db   248 GCCCAGACTCTAGATAGACACGCTTTC------------------CTT 277

Qy   132 PheThrSerAsnProHisAspLeuValProPhePheGluTyrLysLeuTyrAlaAsnGly 151
Db   278 GAACTGAAATTTCCCATAAAGAAAACCAGATTTTGGAGTTCGTTCTT------ 325

Qy   152 AspValProAspLeuSerGlyLeuAlaSerProArgThrPheAlaThrHisLeuProPhe 171
Db   326 ------------GAAATCCTCACCACAACTGAACACAATCTCCCTTCA 367

Qy   172 GlySerLeuLysGlyThrIleGluLysProGlyValValLysValTyrLeuCysArgAsn 191
Db   368 CACTGATTCCACCATCTATCTCGAAAAGAAAACTGCAAGAATGTCTATGTGGCCCGAAAT 427

Qy   192 ProPheAspThrPheIleSerSerTrpHisTyrThrArgAsnAsnIleLysSerGluVal 211
Db   428 CCCAAGGAATTGCCTGGGTCCTACTACACCAGCTTT------ATGAGGAAA 484

Qy   212 SerProValLeuLeuAspAspGlnAlaPheAspLeuTyrCysArgGlyValIleGlyPheGly 231
Db   485 GATCCTCAGAACTTAGACGAAATTTTATGAGAAACTTCATGTCCGGAAAAAGTTGTTGGCGGG 544

Qy   232 ProPheTrpGluHisMetLeuGlyTyrTrpArgGluSerLeuLysArgProGluLysVal 251
Db   545 TCCTGGTTTGACCATGTGAAAGGATGGTGG------GCTGCAAAAGACATGCACCGGATC 598

Qy   252 PhePheLeuArgTyrGluAspLeuLysAspAspIleGluThrAsnLeuLysArgLeuAla 271
Db   599 CTCTACCTCTTCTACGAGGATATTAAAAAAAATCCAAAACATGAGATCCACAAGGTGTTG 658

Qy   272 ThrPheLeuGluLeuProPheThrGluGluGluArgLysGlyValValLysAlaIle 291
Db   659 GAATTCTTGGGAGAAAAACTTGGTCAGGTGAT------GTTATAAACAAGATT 703

Qy   292 AlaGluLeuCysSerPheGluAsnLeuLys 303
Db   704 GTCCACCATACCTCATTTGATGTAATGAAGGATAATCCCATGGCCAACCATACTGCGGTA 763   LysLeu

Qy   304 GluValAsnLysSerAsnLysSerIleLysSerAsnArgPheLeuPheArgLys 323
Db   764 CCTGCTCACATATTCAATCACTCCATCTCAAAATTT------ATGAGGAAA 808

Qy   324 GlyGluValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArgLeuSerAla 343
Db   809 GGGATGGCTGGAGACTGGAAGAACCACTTACGTGGCTTTGAATGAGAAACTTTGATTAAG 868

Qy   344 LeuValAspAspLysLeuGlyLySerGlyLeuThrPheArgLeu 358
Db   869 CATTATGAAAGAAGATGGCAGGGTCCACACTGAACTTCTGCCTG 913

US-10-019-931-3

Qy   86 HisPheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThr 105
Db   128 AATTTCCAAGCCAAGCCTGATGATCTTATTCTGGCAACTTACCCAAAGTCAGGTACAACA 187

Qy   106 TrpIleLysLysAlaLeuThrPheThrIleLeuAsn------------------ 116
Db   188 TGGATGCATGAAATTTTAGACATGATTGTTGAGAAAAATGCAAAAGA 247

Left column

```
FEATURE:
  NAME/KEY: CDS
  LOCATION: (426)...(1308)

US-09-328-174A-2

Alignment Scores:
  Pred. No.:                        1.81e-25        Length:        1396
  Score:                            294.00          Matches:       77
  Percent Similarity:               47.48%          Conservative:  55
  Best Local Similarity:            27.70%          Mismatches:    120
  Query Match:                      15.49%          Indels:        26
                                    3               Gaps:          8

US-10-019-931-3 (1-359) x US-09-328-174A-2 (1-1396)
```

```
Qy   87 PheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThrTrp 106
           ||||||||| |||| |||| ||| |||||||
Db  525 TTCCAGGGCCCGGCCTGATGACCTGCTCATCAGCACCTACCCCAAGTCCGGCACCACCTGG 584

Qy  107 LeuLysValAlaLeuThrProPheThrIleLeuLeuAsnArgHisArgPheAspProValAlaSerSer 126
          :::::|
Db  585 GTGAGGCCAGATTCGGACATGATC------------TACCAGGCGCGGTGACCTGGAAAAGTGT 635

Qy  127 ThrAsnHisProLeuPheThrSerAsnProHisAspLeuValLeuAlaThrIleProLysGluThrTyrLys 146
          |||||||||||||||:|||
Db  636 CACCGAGCTCCCATCTTCATGCGG------GTGCCCTTCCTTGAGTTCAAA 680

Qy  147 LeuTyrAlaAsnGlyAspValProAsp------------LeuSerGlyLeuAlaSerPro 162
          :::
Db  681 GTC----------CCAGGGATTCCCTCAGGGATGGAGACTCTGAAAAACACCAGCCCA 731

Qy  163 ArgThrPheAlaThrHisLeuProPhePheGlySerLeuLysGluThrIleGlyLysProGly 182
          |||       ||:::
Db  732 CGACTCCTGAAGACACACCTGCCCCTGGCCTTCCCAGACTCTGTTGGATCAGAGAG 791

Qy  183 ValLysValValThrTyrLeuCysArgAspAsnProPheAspThrPheIleLeuSerSerTrpHisTyr 202
          ||||||| |||||||
Db  792 GTCAAAGTGGTTCTATGTTGCCGCGACCAAAGGATGGGGCGGGTTCCTACCACTTC 851

Qy  203 ThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspLeuAlaPheAspLeu 222
          |||:::
Db  852 TACCACAATGGCCAAAGTGTACCCT------CACCCTGGGACCCTGGGAAAGCTTCCTGGAAG 908

Qy  223 TyrCysArgGlyValIleGlyPheGlyTrpGluHisMetLeuGlyTyrTrpArg 242
          |||
Db  909 TTCATGGGCTCGGAGAAGTGTCCCTATAGGGTCATGGGCTTCAAGAGTGGGTGG 965

Qy  243 GluSerLeuLysArgProGlyLysValPhePheLeuArgTyrGluLysLeuLysSerAspAsp 262
          |||:::||
Db  966 GGACTGAGCCGCGACCACCCTGTTCTCTRAATGAAGACATGAAGAACATGAAGAAC 1022

Qy  263 IleGluThrAsnLeuLysSerArgLeuAlaThrPheLeuGluLeuProPheThrGluGlu 282
          :::
Db 1023 CCCAAAAGGAGATTCAAAAGATCCTGGAGTTGTGGGGGCCGCTCCCTGCCAGAGGAG--- 1079

Qy  283 GluArgLysGlyValValValLysAlaIleAlaLeuGluLeuCysSerProGluGluAsnLeuLysLys 302
          |||
Db 1080 ------------ACTGTGGACCTCATGGGTTGAGCACGTGCTTGTTCAAGGAGGATGAAGAAG 1127

Qy  303 LeuGluValLeuAsnLeuSerAsnLysSerIleIleLysAsnPheLeuAsnArgPheLeu 320
          :::
Db 1128 AACCCTATGACCAACTACACCACCCGTCCGCGGGAGTTCATGGACCACAGCATCTCCCCC 1187

Qy  321 ---PheArgLysGlyValGluValIleSerAspTrpValAsnTyrLeuSerProSerGlnValGlu 339
          |||
Db 1188 TTCATGAGGAAGGGCATGGCTGAGAAGACCACCTTCACCTGGCCGCGAATGAG 1247

Qy  340 ArgLeuSerAlaLeuValAlaAspLysGlyLeuGlyLysGlySerGlyLeuThrPheArg 357
          |||:::
Db 1248 CGCTTCGATGCCGGACTATGCCGGAGGAGAGAGATGCGCAGGCCTGCCAGCTTCCGC 1301
```

Right column

```
Sequence 4457, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
TITLE OF INVENTION:
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 4457
  LENGTH: 1045
  TYPE: DNA
  ORGANISM: Human
US-09-949-016-4457

US-10-019-931-3 (1-359) x US-09-949-016-4457 (1-1045)

Alignment Scores:
  Pred. No.:                        9.35e-25        Length:        1045
  Score:                            286.50          Matches:       85
  Percent Similarity:               47.33%          Conservative:  57
  Best Local Similarity:            28.33%          Mismatches:    121
  Query Match:                      15.09%          Indels:        37
                                    3               Gaps:          11
```

```
Qy   75 LysGluIleLeuGlnAlaLeuIleMetSerPheGln------------------------Lys 85
          :::||                                                   :::
Db  140 GAAGAAGTCCATGGAGGATTCTAAGTGTATAAAAGATTTGTGGGATAAATGTGGAA 199

Qy   86 HisProGlnSerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThr 105
          |||||||:::||||||| |||||||||||
Db  200 GCGTTCCAGGGAAGACCAGATGATCTTGTCATTGCCACCTACCCTAAATCTGGTTACACC 259

Qy  106 TrpLeuLysValAlaLeuGluThrPheThrIleLeuLeuAsnArgHisArgPheAspProValAlaSer 125
          |||||| :::
Db  260 TGGGTTAGTGAAATTGTGATATGATC----------TATAAAGAGGGTGAGTGGAAAG 310

Qy  126 SerThrAsnHisProLeuPheThrSerAsnProHisAspLeuValProPhePheGluTyr 145
          :::||
Db  311 TCAAAGAAGATGTAATTTTT------------AATCGAATACCTTTCCTGGAATGC 355

Qy  146 LysLeu------TyrAlaAsnGlyAspValProAspLeuSerGlyLeuAlaSerProArg 163
          :::                          :::
Db  356 AGAAAAGAAAAAACCTCATGAAATGGA---GTAAAACAATTAGATGAGATGAATTCCTCAGA 412

Qy  164 ThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGlyLysProGlyVal 183
          ||||||||||:::
Db  413 ATTGTGAAGACTCATTGCCACCTGAACTTCTTCCGCCTCATTTGGAAAAGGAATGT 472

Qy  184 LysValValTyrLeuCysArgAspAsnProPheAspThrPheIleLeuSerSerTrpHisTyrThr 203
          :::||||||
Db  473 AAGATAATCTATCTTTGCCGGAATCCAAAGGATGTGGCGTTTCCTTTATTATTTCTTT 532

Qy  204 AsnAsnIleLysSerGluSerValSerProValLeuLeuAspLeuAlaPheAspLeuTyr 223
          :::|||
Db  533 CTAATGTGGCTGGTGCATCCA------AATCCTGGATCCTTCAGAGTTGTGTGGAAATTC 589

Qy  224 CysArgGlyValIleGlyPheGlyProPheGluHisMetLeuGlyTyrTyrArgGlu 243
          :::||||||
Db  590 ATGCAAGGACAGGTTCCTTATGGTTCCTGGGTATAAACATGTAAAAATCTTTGGGTGG 646

Qy  244 SerLeuLysArgProGlyLysValPhePheLeuArgTyrGluAspLeuLysAspAspIle 263
          :::||                                                   :::GAA
Db  647 AAGGGAAAGAGTCCA------CGTGTACTATTTCTTTTCTACGGAGACCTGAAAGAGGATATC 703
```

RESULT 12
US-09-949-016-4457

RESULT 13

US-08-325-562-1

; Sequence 1, Application US/08325562
; Patent No. 5714594
; GENERAL INFORMATION:
; APPLICANT: Weinshilboum, Richard M.
; APPLICANT: Aksoy, Ibrahim A.
; APPLICANT: Wood, Thomas C.
; TITLE OF INVENTION: cDNA CLONING AND EXPRESSION OF HUMAN
; TITLE OF INVENTION: LIVER ESTROGEN SULFOTRANSFERASE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Schwegman, Lundberg & Woessner P.A.
;   STREET: 3500 IDS Center
;   CITY: Minneapolis
;   STATE: MN
;   COUNTRY: USA
;   ZIP: 55402
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/325,562
;   FILING DATE: 18-OCT-1994
;   CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: Mueting, Ann M.
;   REGISTRATION NUMBER: 33,977
;   REFERENCE/DOCKET NUMBER: 150.131US1
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 612-339-0331
;   TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1063 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 107..989
; US-08-325-562-1

Alignment Scores:

Pred. No.:	1
Score:	286.50
Percent Similarity:	47.33%
Best Local Similarity:	28.33%
Query Match:	15.09%
DB:	9.6e-25

Length:	1063
Matches:	85
Conservative:	57
Mismatches:	121
Indels:	37
Gaps:	11

```
Qy   264 GluThrAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGluGluGlu 283
             :::|||||              |||||||||
Db   704 AGAAAAGAGGTGATAAAATTGATACATTCCTGGAAAGGAAGGCCATCAGAGGAG------ 757

Qy   284 ArgLysGlyValValLysAlaIleAlaGluLeuCysSerPheGluAsnLeuLysLeu 303
                                      :::|||           :::|||
Db   758 ------CTTGTGGACAGGATTATACATCATTCGTTCCAAGAGATGAAAGAC--- 805

Qy   304 GluValAsnLysSerAsnLys------ 317
             ||||                    SerIleLysAsnPheGluAsn
Db   806 ------AATCCATCCACAAATTACACAACACTGCCACACGAAATTATGAACCAGAAATTG 859

Qy   318 ArgPheLeuPheArgLysGlyGluValSerAspTrpValAsnTyrLeuSerProSerGln 337
             ||||||                                          :::
Db   860 TCGCCCCTTCATGAGAAAGGGAATTACAGGAGACTGGAAAAATCACTTTACAGTAGCCCTG 919

Qy   338 ValGluArgLeuSerAlaLeuValAspAspLysLeuGlyGlySerGlyLeuThrPheArg 357
             :::|||                                          |||
Db   920 AATGAAAAATTTGATAAACATTATGAGCAGCAAATGAAGGAATCTACACTGAAGTTTCGA 979
```

US-10-019-931-3 (1-359) x US-08-325-562-1 (1-1063)

```
Qy    75 LysGluIleGlnAlaIleMetSerPheGln------------------------Lys 85
             :::|||||:::                              :::
Db   140 GAAGAGTCCATGGGATTCTAAATGTATAAAAGATTTGTCAAATATTGGGATAAGTGGAA 199

Qy    86 HisPheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThr 105
             |||||||:::||||||||||      :::|||||||||:::
Db   200 GGCTTCCAGGCAAGACCAGATCAATCTTCGTTCCATTGCCACCTAAATCTGGTACAACC 259

Qy   106 TrpLeuLysAlaLeuThrPheThrIleLeuAsnArgHisArgPheGluValAlaSer 125
             :::|||          :::                          :::
Db   260 TGGGTTAGTGAAATTGTGTTATATAGATC--- 310

Qy   126 SerThrAsnHisProLeuPheThrSerAsnProHisAspLeuValProPhePheGluTyr 145
             :::                                  :::|||:::
Db   311 TGCAAAGAAGATGTAATTTTT------------------------AATCGAATACCTTCCTGGAATGC 355

Qy   146 LysLeu------------------TyrAlaAsnGlyAspValProAspLeuSerGlyLeuAlaSerProArg 163
             ||||||                              |||||||||
Db   356 AGAAAAGAAAACCTCATGAAATGGA---GTAAAACAATTAGATGAGAGAAATTCTCCTAGA 412

Qy   164 ThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluGlyProGlyVal 183
             |||                        :::                    |||
Db   413 ATTGTGAAGACTCATTTGCCACCTGAACTTCTTCCTGGCTCATTTGGGAAAAGGATTGT 472

Qy   184 LysValValTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrpHisTyrThr 203
             |||:::||||||||||||:::                :::||||||
Db   473 AAGATAATCTATCTTTGCCGGAATCCATTTGATACTTTCATTTCCTCATGGCACTACAACC 532

Qy   204 AsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnIleAlaPheAspLeuTyr 223
             :::|||                              :::                :::
Db   533 CTAATGGTGGCTCGGTCATCCA---AATCCTGGATCCTTCAGAGTTTGTGGAAAATTC 589

Qy   224 CysArgGlyValIleGlyProPheGlyProProTrpGluHisMetLeuGlyTyrProGlu 243
             :::                              :::|||
Db   590 ATGCAAGGACAGGTTCCTTATGGTTCCTGGTAAAACATGTAAAATCTTGGTCG---GAA 646

Qy   244 SerLeuLysArgProGlyLysValPhePheLeuValArgTyrGluAspLeuLysAspAspIle 263
             |||||||              |||                    :::|||
Db   647 AAGGGAAAGAGTCCA---CGTGTACTATTTCTTTTCTGCAGAGACCTGAAGAGGATATC 703

Qy   264 GluThrAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGluGluGlu 283
             :::                              |||||||
Db   704 AGAAAAGAGGTGATAAAATTGATACATTCCTGGAAAGGAAGGCCATCAGAGGAG------ 757

Qy   284 ArgLysGlyValValLysAlaIleAlaGluLeuCysSerPheGluAsnLeuLysLeu 303
                                          :::|||
Db   758 ------CTTGTGGACAGGATTATACATCATTCGTTCCAAGAGATGAAAGAC--- 805

Qy   304 GluValAsnLysSerAsnLys------ 317
             ||||                    SerIleLysAsnPheGluAsn
Db   806 ------AATCCATCCACAAATTACACAACACTGCCACACGAAATTATGAACCAGAAATTG 859

Qy   318 ArgPheLeuPheArgLysGlyGluValSerAspTrpValAsnTyrLeuSerProSerGln 337
             ||||||                              :::
Db   860 TCGCCCCTTCATGAGAAAGGGAATTACAGGAGACTGGAAAAATCACTTTACAGTAGCCCTG 919

Qy   338 ValGluArgLeuSerAlaLeuValAspAspLysLeuGlyGlySerGlyLeuThrPheArg 357
             :::|||                              |||
Db   920 AATGAAAAATTTGATAAACATTATGAGCAGCAAATGAAGGAATCTACACTGAAGTTTCGA 979
```

RESULT 14

US-08-437-795-1

; Sequence 1, Application US/08437795
; Patent No. 574355
; GENERAL INFORMATION:
; APPLICANT: Weinshilboum, Richard M.
; APPLICANT: Aksoy, Ibrahim A.
; APPLICANT: Wood, Thomas C.
; TITLE OF INVENTION: cDNA CLONING AND EXPRESSION OF HUMAN
; TITLE OF INVENTION: LIVER ESTROGEN SULFOTRANSFERASE

; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg & Woessner P.A.
; STREET: 3500 IDS Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/437,795
; FILING DATE: 09-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueting, Ann M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 150.137US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1063 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..989

US-08-437-795-1

Alignment Scores:
Pred. No.:
Score: 286.50
Percent Similarity: 47.33%
Best Local Similarity: 28.33%
Query Match: 15.09%
DB: 2

Length: 1063
Matches: 85
Conservative: 57
Mismatches: 121
Indels: 37
Gaps: 11

Score 9.6e-25

US-10-019-931-3 (1-359) x US-08-437-795-1 (1-1063)

Qy   75 LysGluIleGlnAlaIleMetSerPheGln----------------------------------Lys 85
        ||:|:||||:|| ::|
Db 140 GAAGAAGTCCATGGGGATTCTTAATGTATATAAAGATTTTGTCAAATATTGGGATAATGTGGAA 199

Qy   86 HisPheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThr 105
        |||||||:||| ||||||||||
Db 200 GCGGTTCCAGGAGCACAAGACAGATACTTGTCATTGCCACCTACCTTAAATCGGTACAACC 259

Qy 106 TrpLeuLysAlaLeuThrPheThrIleLeuAsnArgHisArgProValAlaSer 125
        |:|:   ::|:
Db 260 TGGGTTAGTGAAATTGTGTATATGATC----------TATAAAGAGGGTGATGTGAAAAAG 310

Qy 126 SerThrAsnHisProLeuPheThrSerAsnProHisAspLeuValProPhePheGluTyr 145
        ||||:||||| |||
Db 311 TGCAAAGAAGATGTAATTTTT---------AATCGAATACCTTTCCTGGAATGC 355

Qy 146 LysLeu-----TyrAlaAsnGlyAspValProAspLeuSerGlyLeuAlaSerProArg 163
        |||| ||||
Db 356 AGAAAAGAAAACCTCATGAATGGA----GTAAACAATTAGATGAGATGAATTCTCCTAGA 412

Qy 164 ThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluLysProGlyVal 183
        ||||||||
Db 413 ATTGTGAAGACTCATTTGCCACCTGAACTTCCTTCCTGCCATTTTGGGAAAAGATTGT 472

Qy 184 LysSerValValTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrpProHisTyrThr 203
        ||:|:||||:|||
Db 473 AAGATAATCTATCTTTGCCGAATGCAAAGGATGTGGCGTGTTCCTTTTATTATATTTCTTT 532

US-09-717-321A-45

; RESULT 15
; US-09-717-321A-45
; Sequence 45, Application US/09717321A
; Patent No. 6852845
; GENERAL INFORMATION:
; APPLICANT: Gould-Rothberg, Bonnie E.
; APPLICANT: DiPippo, Vincent A.
; APPLICANT: Ramseh, Tennore M.
; APPLICANT: Gerwein, Robert W.
; TITLE OF INVENTION: METHOD OF IDENTIFYING TOXIC AGENTS USING NSAID-INDUCED
; TITLE OF INVENTION: DIFFERENTIAL GENE EXPRESSION IN LIVER
; FILE REFERENCE: 15966-601 Utility
; CURRENT APPLICATION NUMBER: US/09/717,321A
; CURRENT FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/166,923
; PRIOR FILING DATE: 1999-11-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 45
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Rattus norvegicus

US-09-717-321A-45

Qy 204 AsnAsnIleLysSerGluSerValSerProValSerLeuLeuAsnGlnAlaPheAspLeuTyr 223
        ::||||
Db 533 CTAATGGTGGCTGGGTCGTCATCCA----AATCCTGGATCCTTTCCAGAGTTTGTGGAGAAATTC 589

Qy 224 CysArgGlyValIleGlyPheGlyProPheTrpGluHisMetLeuGlyTyrTrpArgGlu 243
        ::|| ||||:::|||
Db 590 ATGCAAGGACACAGGTTCCTTATGGTTCCTGGGTATAAACATGTAAAATCTTGGTGG----GAA 646

Qy 244 SerLeuLysArgProGlyLysValPhePheLeuArgThrGluAspLeuLysAspAspApIle 263
        ||||||||| :||||::||||
Db 647 AAGGGAAAGAGTCCA----CGTGTACTATTTCTTTTCTACGAAGACCTGAAAGAGGATATC 703

Qy 264 GluThrAsnLeuLysArgLeuArgHisArgPheLeuAlaThrPheLeuGluLeuProPheGluGluGlu 283
        :::| ::|||||||||||
Db 704 AGAAAAGAGGTGATAAAATTGATACATTTTCCTGGAAAGGAAGCCATCAGAGGAG---- 757

Qy 284 ArgLeuGlyValValLysAlaIleAlaGluLeuCysSerPheGluAsnLeuLysLysLeu 303
        |||||||||
Db 758 ----------CTTGTGGACAGGAATTATACATCATTTCCAAGAGATGAAGAAC---- 805

Qy 304 GluValAsnLysSerAsnLys--------------------SerIleLysAsnPheGluAsn 317
        ::|:   :::
Db 806 ------AATCCATCCACAAATTACACAACACTGCCAGACGAGAAATTATGA 859

Qy 318 ArgPheLeuPheArgLysGlyGluValSerArgAspTrpValAsnTyrLeuSerProSerGln 337
        ||||||||| :::|
Db 860 TCGCCCTTCATGAGAAAGGAATTCACCTTTACAGTAGCCCTG 919

Qy 338 ValGluArgLeuSerAlaLeuValAspAspLysSerLeuGlyGlySerGlyLeuThrPheArg 357
        :::| |||
Db 920 AATGAAAAATTTGATAAACATTATGGACAGCAGAAATGAGAGGAATTCTACACTGAAGTTTCGA 979

US-10-019-931-3 (1-359) x US-09-717-321A-45 (1-993)

Alignment Scores:
Pred. No.:
Score: 278.50
Percent Similarity: 47.12%
Best Local Similarity: 27.56%
Query Match: 14.67%
DB: 3

Length: 993
Matches: 86
Conservative: 61
Mismatches: 121
Indels: 45
Gaps: 14

Score 8.09e-24

US-10-019-931-3 (1-359) x US-09-717-321A-45 (1-993)

Qy   57 ArgGlyTyrTrpArgThrArgTyrLeuTyrLeu-------------------- 66
        :::|||||| ||||:::|:
Db   4 AAGGGCTGGAATACTAAAAG-TTATTCATGATGTCAGACTATACTTGGTTGAAGGAATA 62

Qy   67 ---PheGlnGlyPheTrpCysGlnAlaLysGlyIleGlnAlaIleMetSerPheGlnLys 85
        |||
Db   85

```
Db   63  CCTTTTCCTGCCTTTTGGTTT-----TCCAAGAAATT-------CTGAAAAATAGTTGTAAG 113

Qy   86  HisPheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThr 105
         |||   |:::|||||
Db  114  AAGTTTGTGGTAAAAGAGGACGACTTGATCATATTGACTTACCCCAAGTCAGGAACGGAAC 173

Qy  106  TrpLeuLysAlaLeuThrPheThrIleLeuAsnArgHisArgPheAspPro------Val 123
         ||||||||       :::
Db  174  TGGCTG------ATCGAGATTGTCTCGCTTGATTCAGACCAAGGGAGGATCCCAAGTGGATC 227

Qy  124  AlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAspLeuValProPhePhe 143
                                                     |||||::::
Db  228  CAATCTATG------CCCATCTGG--------GATCGGCTCACCCTGGATA 263

Qy  144  GluTyrLysLeuTyrAlaAsnGlyAspValAspLeuSerGlyLeuAlaSerProArg 163
         |||||:|       |||||||||
Db  264  GAG------ACTGGTTCAGGATATGATAAATTAACAAAAATGGAAGGACCACGA 311

Qy  164  ThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluLysProGlyVal 183
         :::|||||||||
Db  312  CTCATGACCTCCATCTTCCATCGATCTTTTCTCCAAGTTCTCTTCAGTTCCAAGGCC 371

Qy  184  LysValValTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrpHisTyrThr 203
         ||||||||||||       ::::::::|||
Db  372  AAGGTGATATATCTCATCAGAAATCCCAGAGATGTTCTTGTTTCTCCTTATTTTTCTGG 431

Qy  204  AsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaPheAspLeuTyr 223
         ::::|||||||||||
Db  432  AGTAAGATCGCCCTGGAGAAG---AAACCAGACTCGGTGGGAACTTACGTTGAATGGTTC 488

Qy  224  CysArgGlyValIleGlyPheGlyProPheTrpGluHisMetLeuGlyTyrTyrTrpArgGlu 243
         :::|||||||||||||       :::|||||
Db  489  CTCAAAGGAGAAATGTTCGATATGGATCATGGGTTTGAGCACATCCGTGGCTGGCTG---- 542

Qy  244  SerLeuLeuArgProGlyLysValPhePheLeuArgThrTyrGluValLysAspAspIle 263
         |||||||||:::|||
Db  543  TCTATGAGAGAATGGGACACTTCTTGGGTACTGTACTATGAAGACATGAAAAAGGATACA 602

Qy  264  GluThrAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGluGluGlu 283
         :::::|||||||||
Db  603  ATGGGGATCCATAAAGAAGATAATGTGACTTCCTCGGGAAAAAAATTAGAGCCAGATGAG--- 659

Qy  284  ArgLysGlyValValLysAlaIleAlaGluLeuCysSerPheGluAsnLeuLysLysLeu 303
         :::       ||||||||:::|||
Db  660  --------------CTGAATTGGTCCTCCAAGTATAGTTCCTTCCAAGTCGGTGAAA----- 701

Qy  304  GluValAsnLysSerAsnLysSerIleLeuLysAsnPheGlu--------AsnArgPheLeu 320
         |||       :::
Db  702  GAAAACAACATGTCCAATTATAGCCCTCATGGAGAAGGAACTGATTCTTACTGGTTTTACT 761

Qy  321  Phe---ArgLysGlyGluValSerAspTrpValAsnTyrLeuSerProSerGlnValGlu 339
         |||   |||||||||
Db  762  TTCACTGAGAAAAGGCACAACTAATGACTGGAAGAATCACTTCACAGTAGCCCAAGCTGAA 821

Qy  340  ArgLeuSerAlaLeuValAlaAspAspLysLeuGlyGly 351
         :::||||||||
Db  822  GCCTTTGATAAAGTGTTCCAGGAGAAAATGGCGGGT 857
```

OM protein - nucleic search, using frame_plus_p2n model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: December 23, 2005, 19:08:24 ; Search time 834 Seconds
(without alignments)
3559.600 Million cell updates/sec

Title:
Perfect score: 1898
Sequence: 1 MATSSMKSIPMAIPFSMCH..........RLSALVDDKLGGSGLTFRLS 359

Searched: 9793542 seqs, 413469005 residues

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Total number of hits satisfying chosen parameters: 19587084

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10019931/runat_16122005_165220_18500/app_query.fasta_1.519
-DB=Published_Applications_NA_Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10019931@CGN_1_1549 @runat_16122005_165220_18500 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1058.5	55.8	1311	7	US-10-424-599-58871	Sequence 58871, A
2	770.5	40.6	996	3	US-09-938-842A-1027	Sequence 1027, Ap
3	770.5	40.6	996	3	US-09-938-842A-1027	Sequence 1027, Ap
4	763.5	40.2	1069	8	US-11-767-795-558	Sequence 558, App
5	757	39.9	981	3	US-09-938-842A-2049	Sequence 2049, Ap
6	757	39.9	981	3	US-09-938-842A-2049	Sequence 2049, Ap
7	725.5	38.2	787	7	US-10-424-599-82242	Sequence 82242, A
8	674	35.5	1041	7	US-10-437-963-55857	Sequence 55857, A
9	672	35.4	1440	7	US-10-767-795-1749	Sequence 1749, Ap
10 c	671.5	35.4	271990	6	US-10-195-144-87	Sequence 87, Appl
11 c	671.5	35.4	271990	6	US-10-345-072-87	Sequence 87, Appl
12	666.5	35.1	1251	8	US-10-767-795-1750	Sequence 1750, Ap
13	666	35.1	1358	7	US-10-739-930-1663	Sequence 1663, Ap
14	659.5	34.7	1077	7	US-10-437-963-55524	Sequence 55524, A
15	648.5	34.2	1579	8	US-10-425-115-167844	Sequence 167844, A
16	635	33.5	1021	7	US-10-437-963-94656	Sequence 94656, A
17	633	33.4	1068	7	US-10-437-963-73855	Sequence 73855, A
18	614	32.3	1050	6	US-10-259-194A-15	Sequence 15, Appl
19	614	32.3	1050	6	US-10-437-963-54999	Sequence 54999, A
20	613.5	32.3	2067	6	US-10-195-144-80	Sequence 80, Appl
21	613.5	32.3	2067	6	US-10-345-072-80	Sequence 80, Appl
22	610	32.1	1410	8	US-10-425-115-101459	Sequence 101459,
23	609	32.1	1366	7	US-10-767-701-10838	Sequence 10838, A
24	601.5	31.7	1011	7	US-10-437-963-96500	Sequence 96500, A
25	598	31.5	1017	6	US-10-259-194A-3	Sequence 3, Appli
26	598	31.5	1017	7	US-10-437-963-27879	Sequence 27879, A
27	595.5	31.4	1056	7	US-10-437-963-51349	Sequence 51349, A
28	592	31.2	1403	7	US-10-425-114-21478	Sequence 21478, A
29	592	31.2	1526	7	US-10-425-114-27795	Sequence 27795, A
30	591.5	31.2	1461	6	US-10-259-194A-131	Sequence 131, App
31	591.5	31.2	1461	7	US-10-260-228-1003	Sequence 1003, Ap
32	589.5	31.1	1482	7	US-10-437-963-75177	Sequence 75177, A
33	589	31.0	984	6	US-10-259-165-443	Sequence 443, App
34	589	31.0	1435	7	US-10-437-963-101470	Sequence 101470,
35	589	31.0	1032	6	US-10-259-194A-329	Sequence 329, App
36	588.5	31.0	1355	7	US-10-437-963-56891	Sequence 56891, A
37	588.5	31.0	1041	6	US-10-259-165-331	Sequence 331, App
38	586.5	30.9	1041	7	US-10-437-963-68828	Sequence 68828, A
39	583	30.7	1828	8	US-10-425-115-168303	Sequence 168303,
40	550.5	29.0	1192	3	US-09-854-122-15	Sequence 15, Appl
41	546	28.8	1033	7	US-10-437-963-72606	Sequence 72606, A
42	545.5	28.7	1435	7	US-10-437-963-101470	Sequence 101470,
43	529.5	27.9	1044	7	US-10-437-963-49158	Sequence 49158, A
44	521.5	27.5	1878	7	US-10-437-963-72617	Sequence 72617, A
45	520.5	27.4	1019	7	US-10-767-701-15199	Sequence 15199, A

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
US-10-424-599-58871
Sequence 58871, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 58871
LENGTH: 1311
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(1311)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_24172C.1
US-10-424-599-58871

Alignment Scores:
Pred. No.:   5.22e-119   Length: 1311
Score: 1058.50   Matches: 191
Percent Similarity: 74.71%   Conservative: 66

Best Local Similarity: 55.52%
Query Match: 55.77%
DB:

		Mismatches:	83
		Indels:	4
		Gaps:	7

US-10-019-931-3 (1-359) x US-10-424-599-58871 (1-1311)

Qy	19	CysHisLysLeuLeuLeuLys-GluGlyLysThrArgAspValProLysAlaGluGl	38
		|||||||||||||||||: ::	
Db	53	TGTCACATGCTTCAGAGAAGAAAATGAATCCGAGAAAGGGAGGAAATAACAATAGGAGA	112

Qy	38	uAspGluGlyGlyLeuSerCysGluPheGlnGluMetLeuAspSerLeuProLysGluArgGl	58
		|||||||||: : :::	
Db	113	AGACAAAG---CTAAGTCAAGAATGTAAGAGTTGATACTCTCTTCCTTCCTAGGGAGGAGG	169

Qy	58	YIrpArgThrArgTyrLeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGluIleGl	78
		||||||||||:|||||||:||||||||	
Db	170	TTGGAGAACACGGTTATATATATTCTATTTCCAAGGATTTTGGTGCCAGCAGCAGAAATCCA	229

Qy	78	nAlaIleMetSerPheGlnHisPheGlnSerLeuGluAsnAspValValLeuAlaTh	98
		||||||||||||:::||||||||:||||||:::||||||	
Db	230	AGCAATAATCACTTTTCAGAGAGCACTTCCCAAGCTAAAGACAGTAGAGTTATTGTGGCCAC	289

Qy	98	rIleProLysSerGlyThrArgThrTrpLeuLysProPheThrThrIleLeuAsnArgHi	118
		||||||||||||||||:||||:|||||||||:|||||||||:|||||	
Db	290	AATTCCAAAATCAGGTACCACTTGGCTAAAGCCTTCACCTTTGCCATTGTTCAATCGCCA	349

Qy	118	sArgPheAspProValAlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAs	138
		||| |||||||:|||||:|||||:||||:|||||||	
Db	350	TACTCCATAGTATCACTACACTAGACAAGTCACATCCACTTTGCTTACTTCACTTACTCATGA	409

Qy	138	pLeuValProPheGluTyrLysLeuTyrAlaAsn------GlyAspValProAspLe	156
		||||||||||||:||||||||||||| :::||||||:::|||	
Db	410	ACTTGTGCCCTTTCATAGAATACACCGTTTATGGTAATGCCCCTAGCCATGTTCCAAACCT	469

Qy	156	uSerGlyLeuAlaSerProArgThrProPheAlaThrHisLeuProPheGlySerLeuLysGl	176
		||:|| ::||||||||||||||||:::||||||	
Db	470	ATCCAACATGACTGAGCCACCAGACACTTTTGTACACATATTCCATTCCATTTGGCCAA	529

Qy	176	uThrIleGluGlyValProGlyValValSerValThrLeuTyrLeuArgAsnProPheAspThrPh	196
Db	530	GTCCATCAAGGAGTCCAAATAGTAGAATAATTATATATGTAGGAACCCACTTGACACTTT	589

Qy	196	eIleSerSerTrpHisTyrThrArgAsnAsnIleLysSerGluSerValSerProValLeuLe	216
		|||:||||||||||||:||||:|||:|||| ::	
Db	590	TGTGTCTCACTTGGATTTCCTCAACAAAATTAAGCCCAGAACATTTACCTGAATTTGAAACT	649

Qy	216	uAspGlnAlaPheAspLeuuTyrCysArgGlyValAlaIleGluGlyPheGlyProPheTrpGluHi	236
		::||||||| |:|||:|	
Db	650	AGGGGAGCTTTTGAAAAGTATTGCAAAGGAATAATAAGGGTTTTGGTCCAACTTGGGACCA	709

Qy	236	sMetLeuGlyTyrIrpArgGluSerLeuLysArgProGlyLysValPhePheLeuArgTy	256
Db	710	AATGTTGGGTTATTGGAAAGGAGAGTAGCCTAGGAGGTTTGTTCTTTGAAGTA	769

Qy	256	rGluAspProLeuLysAspAspAsnIleGluThrArgLeuuAlaThrPheLeuGluLe	276
		||||||||||:::||| :::	
Db	770	CGAGGAGATCTTAAAAAGAGTCAATTTCATGTGAAAAGATAGCGGAGTTCTTAGGATG	829

Qy	276	uProGluThrGluGluGluGluuArgGlyLysGlyValAlaIleAlaGluLeuCysSe	296
Db	830	TCCTTTCACTTCGGAGGAAGAAGGTGGACTATTGAGAGCATAATCAAGCTATGCAG	889

Qy	296	rPheGluAsnLeuLysLeuGluValAlaAsnLysSerIleLysAsnPheGl	316
		:::	
Db	890	CTTCGAGAAGATGAAGGAATTGGGAGGCCAAAAAATCTGGAACATTTGCTAGGAACTTTGA	949

Qy	316	uAsnArgPheLeuuPheArgLysGlyValAlaSerAspTrpValAlaAsnTyrLeuSerProSe	336
		::||||||||||:|||||| :||:|	
Db	950	GAGAAAAGTACTTGTTCCGAAAGGCTGAAAAATGGGAGACTACCTTTCCCCTGA	1009

Qy	336	rGlnValGlyArgLeuSerAlaLeuValLysAspAspLysLeuGlyGlySerGlyTyrLeuThrPh	356
		|||||||:||||:: ||||||||||||||||||	
Db	1010	AATGGGGTGAAAAGTTATCGCAAAATTATGGGAGAAAAAGTTAAGTGTGGGTCAGGCCTTGTCATT	1069

Qy	356	eArgLeuSer	359
		:::|||||	
Db	1070	TTAAGTGTCT	1079

RESULT 2
US-09-938-842A-1027
; Sequence 1027, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1027
; LENGTH: 996
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1027

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

Alignment Scores:		
Pred. No.:	8.94e-84	
Score:	770.50	
Percent Similarity:	62.50%	
Best Local Similarity:	41.67%	
Query Match:	40.60%	

	Length:	996
	Matches:	140
	Conservative:	70
	Mismatches:	121
	Indels:	5
	Gaps:	3

US-10-019-931-3 (1-359) x US-09-938-842A-1027 (1-996)

Qy	22	LeuGluLeuLeuLysGluGlyLysThrArgAspValProLysAlaGluGluAspGluGly	41
		|:|||||||	
Db	1	ATGGAGGCTTCTAAAGAA------GCTCATCACCTTCCAAACTACATGAAAGACGACAAC	54

Qy	42	LeuuSerCysGluPheGlnGluMetLeuAspSerLeuProLysGluArgGlyTyrTrpArgThr	61
		::||||||:|||: : :::	
Db	55	GTTAGTCAAGAAAACCAAGAAACTTGACACTCTTCTACCTTCAGACAAAGATTTCATGGGT	114

Qy	62	ArgTyrLeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGluIleGlnAlaIleMet	81
		|||||||||||:::|||||| :::	
Db	115	TATGGTCGTCTACAACTACAAGGTTGTTGGTACTATCCAAAACACCCAAGCCGGTTCTT	174

Qy	82	SerPheGlnHisPheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLys	101
		|||||||:||:|| :::	
Db	175	GACGTCCAAAAAACACTTCAGCCACGAGATACTGATATAATCCTCGGCTTCTTTGCCCAAA	234

Qy	102	SerGlyThrArgThrTrpLeuLysProPheThrThrIleLeuAsnArgHisArgPheAsp	121
		::||||||:||	
Db	235	GGTGGAACCACTTGGGTCAAATCCCTAAATTTTCGCGTTGTACATGAGAGAAAGTACCGC	294

Qy	122	ProValAlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAspLeuuValPro	141
		:::	
Db	295	------GGAACCCCTCAAACACACATCCTTTGCTCTACAAAACCCTTACAACCCGGTTGTT	348

Qy	142	PhePheGluTyrLysLeuuTyrLeuTyrAlaAsnGlyAspValProAspLeuuSerGlyLeuAlaSer	161
		||| :::|||	
Db	349	TTTCTTGAGGTTGAGTTATACGCTAATAGCCAAATACCGGATCTCGGCAAAAGTATTCTTCT	408

Qy	162	ProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluLysPro	181
		|||::||||||| :::	
Db	409	CCTATGATCTTTTCTACACACATTACACTTACAAGCAATTGCGTGAAGCCACACACAAAA---	465

US-10-019-931-3 (1-359) x US-10-424-599-58871 (1-1311)

Qy	356	eArgLeuSer	359
		:::|||||	
Db	1070	TTAAGTGTCT	1079

```
RESULT 3
US-09-938-842A-1027
Sequence 1027, Application US/0993842A
Publication No. US20040009476A9

GENERAL INFORMATION:
 APPLICANT: Harper, Jeff
 APPLICANT: Kreps, Joel
 APPLICANT: Wang, Xun
 APPLICANT: Zhu, Tong
 TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 TITLE OF INVENTION: SAME, AND METHODS OF USE
 FILE REFERENCE: SCRIP1300-3
 CURRENT APPLICATION NUMBER: US/09/938,842A
 CURRENT FILING DATE: 2001-08-24
 PRIOR APPLICATION NUMBER: US 60/227,866
 PRIOR FILING DATE: 2000-08-24
 PRIOR APPLICATION NUMBER: US 60/264,647
 PRIOR FILING DATE: 2001-01-16
 PRIOR APPLICATION NUMBER: US 60/300,111
 PRIOR FILING DATE: 2001-06-22
 NUMBER OF SEQ ID NOS: 5379
 SEQ ID NO 1027
 LENGTH: 996
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana
US-09-938-842A-1027
```

```
Qy  182 GlyValLysValValThrTyrLeuCysArgAsnProPheAspThrThrPheIleSerSerTrpHis 201
Db  466 GCTTGCAAAACCGTATATGTGTAGAGGTATCAAAGATAACGTTTTGTCCCGGTGGCAT     525

Qy  202 TyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaPheAsp     221
Db  526 TATAGAAACAATGTTCATCCGACAGATGGATCAAGCCACTTTGACTCATGTTGAT         585

Qy  222 LeuTyrCysArgGlyValValIleGlyTyrProPheTrpGluHisMetLeuGlyThrTyrTrp  241
Db  586 GCTTATTGTAGAGGAGTTCTCTTATATGGGACCTTATTGGTACATGTTGGT             645

Qy  242 ArgGluSerLeuLysValPheGlnAlaProGluLysValPheGluAspLeuLysAsp        261
Db  646 AAAGGAGCTTGGAAGCAAAGGACGAAATGTTCTTTTCATGAAGTTACGGAAGATAATTGAG   705

Qy  262 AspIleGluThrArgAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGluGlu   281
Db  706 GAGCCTCGTGTTCAAGTCAACAGAGACTCCGCCGAGTTCTTAGATATTCGGAAATGCCAAGGAA 765

Qy  282 GluGlnArgLysGlyValValLysValAlaIleLeuAlaGluLeuCysSerPheGluAsnLeuLys 301
Db  766 GAAGAGAAAGGTGGATCGGTGAAGGTTGTAGTTTACGGAGAAATTTAAGC              825

Qy  302 LysLeuGluValValAlaSerAsnLysSerIleLeuLysAsnArgPheLeuPhe          321
Db  826 AATTTGGAGGTTAATAAGAATGGACAAGGTTCTTTCATGAAGTTCTTT               885

Qy  322 ArgLysGlyGluValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArgLeu     341
Db  886 AGGAAAGGTGAAGTTGGATGATTGGGAAGAATCATCTTACGGCAACTTTGACTT         945

Qy  342 SerAlaLeuValAspAspLysLeuGlyGlySerGlyLeuThrPheArg                357
Db  946 GATGAGATTATTGACTATAGACTAGGAGACTCCGGTTTGATATTTCAA               993
```

```
Alignment Scores:

Pred. No. :                                   Length:      996
Score:                   770.50              Matches:     140
Percent Similarity:      62.50%              Conservative: 70
Best Local Similarity:   41.67%              Mismatches:  121
Query Match:             40.60%              Indels:        5
DB:                         3                Gaps:          3

US-10-019-931-3 (1-359)  x  US-09-938-842A-1027 (1-996)
```

```
RESULT 4
US-10-767-795-558
Sequence 558, Application US/10767795
Publication No. US20040181830A1
```

```
Qy   22 LeuGluLeuLeuLysGluGluGlyLysThrArgAspValProLysProLysAlaGluGluAspGluGly 41
Db    1 ATGGAGCTTCTAAAGGAA------GCTCATCACCTTCCAAACTACACAGAAGACGAAC       54

Qy   42 LeuSerCysGluPheGlnGluMetLeuAspSerLeuProLysGluArgGlyThr           61
Db   55 GTTAGTTGCAAGAAACCAAGAACTTGATCACTTCTCTACCTTCCAGAAAGATTTCAAGGGT   114

Qy   62 ArgTyrLeuTyrLeuPheGlyGlyPheThrCysGlnIleAlaLysGluIleGlnAlaIleMet   81
Db  115 TATGGTCTCTACAACTACAAAGGTTGTTGGTACATGTTGGT                       174

Qy   82 SerPheGlnLysHisPheGluSerLeuGluLeuAsnAspValValLeuAlaThrIleProLys   101
Db  175 GACGTCCAAAAACACTTCAGGCCACCGAGTACTGATATAATCCTTGCCAAA             234

Qy  102 SerGlyThrThrTrpLeuLysGlyAlaLeuLeuThrPheThrIleLeuLeuMetArgHisArgPheAsp 121
Db  235 GGTGGAACCACTTGGCTCAAATCCCTAATTCTGGTTGTACATAGGCAC               294

Qy  122 ProValAlaSerSerThrArgAsnHisSerProLeuPheThrSerAsnProHisAspAspLeuValPro 141
Db  295 GGAACCCTCAAACACATCCTTTGCTCTTACAAAACCTCATGACCTTGTCCCA           348

Qy  142 PhePheGluTyrTyrLysLeuTyrAlaAlaAsnGlyAspValProAspLeuSerGlyGlyLeuAlaSer 161
Db  349 TTTCTTGAGTTTTGAGTTATACGGCTAATAAGGTACCGAAATCTCCAAAGTATTCTCT     408

Qy  162 ProArgThrPheAlaThrHisIleLeuProPheGlySerLeuLysGluThrIleGluGlyLysPro   181
Db  409 CCTATGATCTTTCTACAGACATGCACTTACAAGCACCACAAA--                   465

Qy  182 GlyValLysValValThrTyrLeuCysArgAsnProPheAspThrThrPheIleSerSerTrpHis  201
Db  466 GCTTGCAAAACCGTATATGTGTAGAGGTATCAAAGATAACGTTTTGTCCCGGTGGCAT     525

Qy  202 TyrThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaPheAsp     221
Db  526 TATAGAAACAATGTTCATCCGACAGATGGATCAAGCCACTTTGACTCATGTTGAT         585

Qy  222 LeuTyrCysArgGlyValValIleGlyTyrProPheTrpGluHisMetLeuGlyThrTyrTrp  241
Db  586 GCTTATTGTAGAGGAGTTCTCTTATATGGGACCTTATTGGTACATGTTGGT             645

Qy  242 ArgGluSerLeuLysValPheGlnAlaProGluLysValPheGluAspLeuLysAsp        261
Db  646 AAAGGAGCTTGGAAGCAAAGGACGAAATGTTCTTTTCATGAAGTTACGGAAGATAATTGAG   705

Qy  262 AspIleGluThrArgAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGluGlu   281
Db  706 GAGCCTCGTGTTCAAGTCAACAGAGACTCCGCCGAGTTCTTAGATATTCGGAAATGCCAAGGAA 765

Qy  282 GluGlnArgLysGlyValValLysValAlaIleLeuAlaGluLeuCysSerPheGluAsnLeuLys 301
Db  766 GAAGAGAAAGGTGGATCGGTGAAGGTTGTAGTTTACGGAGAAATTTAAGC              825

Qy  302 LysLeuGluValValAlaSerAsnLysSerIleLeuLysAsnArgPheLeuPhe          321
Db  826 AATTTGGAGGTTAATAAGAATGGACAAGGTTCTTTCATGAAGTTCTTT               885

Qy  322 ArgLysGlyGluValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArgLeu     341
Db  886 AGGAAAGGTGAAGTTGGATGATTGGGAAGAATCATCTTACGGCAACTTTGACTT         945

Qy  342 SerAlaLeuValAspAspLysLeuGlyGlySerGlyLeuThrPheArg                357
Db  946 GATGAGATTATTGACTATAGACTAGGAGACTCCGGTTTGATATTTCAA               993
```

GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Cao, Yongwei
APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(53534)B
CURRENT APPLICATION NUMBER: US/10/767,795
CURRENT FILING DATE: 2004-01-30
NUMBER OF SEQ ID NOS: 117596
SEQ ID NO 558
LENGTH: 1069
TYPE: DNA
ORGANISM: Gossypium hirsutum
FEATURE:
OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C14123_1
US-10-767-795-558

Alignment Scores:
Pred. No.:
Score: 7.28e-83
Percent Similarity: 64.17%
Best Local Similarity: 43.30%
Query Match: 8

Score:	763.50	Length:	1069
		Matches:	139
		Conservative:	67
		Mismatches:	110
		Indels:	5
		Gaps:	2

US-10-019-931-3 (1-359) x US-10-767-795-558 (1-1069)

```
Qy    12  AlaIleProSerPheSerMetCysHisGlyLeuGluLeuLeuLysGluGlyLysThrArg  31
              ||| ||| :::           :::|:||
Db   119  TCATTACCAGCTGAAAAATGATTGCCAT--------ACTCAAGCTGAAGATTCAGTTTCT 169

Qy    32  AspValProLysAlaGluGluAspGlyLeuLeuSerCysGlyLeuMetLeuAsp  51
          |||||||||||||||||:::|||      ||:::||||||||||||:::||||||||
Db   170  GATGTTCCAAAATACCTGCTAGAATTGGGTCTTCACCCCAGAATGCAAGCAGCTTATATCC 229

Qy    52  SerLeuProLysGlyArgGlyTyrArgThrArgTyrLeuPheGlnGlyPheTrp  71
          |||||||  ||||:::|||   |||||:::|||||||||||||||||
Db   230  TCACTGCCAGCCGAGAAAAGGTTGGGTTGCCAATCTCCTCCAGAATATCAAGCAGGGGTTTGG 289

Qy    72  CysGlnAlaLysGlyGluIleGlnIleAlaIleMetSerPheGlnGlnSerLeuGlu  91
          |||||||:::|||||||:::||||||  :::|||||||:::||||||:::|||
Db   290  CACACCACTCGCCAGTTACAGGCAGTGCTTACTTGCCAAAACCACTTTCAGCCTCAGGAG 349

Qy    92  AsnAspValValLeuAlaThrIleProLysSerGlyThrThrTrpLeuLysAlaLeuThr  111
          |||:::|||   :::|||   |||:::|||
Db   350  ACAGATATCCTCCTTGTTACCACTCCAAAATCAGGAACAACATGGTTAAAGGCTATTGTT 409

Qy   112  PheThrIleLeuAsnArgHisArgPheAspProAlaValAlaSerSerThrAsnHisProLeu  131
          :::|||||||                        :::
Db   410  TTTGCCTTGATCAAACCGAGTCAAGTATCCC---AACACTGATAACAATCACCCTTTG 463

Qy   132  PheThrSerAsnProHisAspLeuValProPhePheGluTyrLysLeuTyrLeuAlaAsnGly  151
          ||||||:::|||||:::|||||||||||||||:::
Db   464  CTCTCAGAGACCCTCACATTCTTGTTCCATTTTGGGCTTTACATTGATAGT 523

Qy   152  AspValProAspLeuSerGlyLeuLeuSerProArgThrPheAlaThrHisLeuProPhe  171
           ||||||||||||||||||||||||
Db   524  CAGGTTCCTAACTTCACCACTTTACATCTCCACGGCTATTCGGACTCATTACCCCTT 583

Qy   172  GlySerLeuLysGluThrIleGluLysProGlyValLysValValLysCysArgAsn  191
          ||||||:::|||:::|||||:::|||||||:::|||||||||||||||:::||||||
Db   584  GTTTCATTGCCAGAATCGGCAAAGAACTCATCTGCAAGCTTGTTATTTATTATGTAGGAAC 643

Qy   192  ProPheAspThrPhePheIleSerSerTrpHisTyrThrAsnAsnIleLysSerGluSerVal  211
          |||||||:::||||||||||||:::||||||||||||:::||||||||||||:::
Db   644  CCAAAGGATACTTTCCTGTATCCGCTTCACTCCAGACAAGCTGAGAACCAAGGATATG 703

Qy   212  SerProValLeuLeuAsnGlnAlaPheAspLeuTyrCysArgGlyValIleGlyTyrPheGly  231
          ||||||:::|||||||:::||||||:::|||
Db   704  GGAAGCAAACTCCTTCGAAGAGACCTTCGATAAGTTACCACAGGGCGTTGGAGTGTATAGGA 763

Qy   232  ProPheTrpGluHisMetLeuGlyTyrTyrTrpArgGluSerLeuLysArgProGluLysVal  251
```

US-09-938-842A-2049

RESULT 5
Sequence 2049, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SCRIP1300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 2049
LENGTH: 981
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-2049

Alignment Scores:
Pred. No.:
Score: 3.99e-82
Percent Similarity: 61.11%
Best Local Similarity: 45.68%
Query Match: 3

Score:	757.00	Length:	981
		Matches:	148
		Conservative:	50
		Mismatches:	120
		Indels:	6
		Gaps:	3

US-10-019-931-3 (1-359) x US-09-938-842A-2049 (1-981)

```
Qy    33  ValProLysAlaGluGluAspGlyLeuLeuSerCysGlyLeuPheGlnGlyLeuMetLeuAspSer  52
          ||||||||:::||||||||||:::|||
Db    19  GTTCCTCGCTTACTTGGGAGATGAAGATCTGACACAAGAACAAGAGCTTCGATCTCTTCT 78

Qy    53  LeuProLysGlyArgGlyTyrArgThrArgTyrLeuPheGlnGlyPheTrpCys  72
          |||||||:::||||||||||:::||||||||||||||||
Db    73  CTTCCTAAAGAAAGGTTAGTGAGTGAAATATATGAAATTCCAAGGACTTTGCAC 138

Qy    73  GlnAlaLysGlyGluIleGlnIleAlaIleMetSerPheGlnGlnSerLeuGluAsn  92
          ||||||:::||||||:::|||||||
Db   139  ACACAACGTATTTTACAAGGAATCTTGATCTGCCAAAAACGCTTTGAGCTAAGATTCC 198

Qy    93  AspValValLeuAlaThrIleProLysSerGlyThrThrTrpLeuLysAlaLeuThrPhe  112
```

RESULT 6
US-09-938-842A-2049
; Sequence 2049, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO: 2049
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2049

Alignment Scores:
Pred. No.:
Score: 757.00
Percent Similarity: 61.11%
Best Local Similarity: 45.68%
Query Match: 39.88%
DB:

US-10-019-931-3 (1-359) x US-09-938-842A-2049 (1-981)

Length: 981
Matches: 148
Conservative: 50
Mismatches: 120
Indels: 6
Gaps: 3

```
Db   781  AAGTGTGTGTGAGCTTTGAGAGTTTAAGTAGTTAATCAAGAAGGAAAATTGCCA   840
Qy   313  LysAsnPheGluAsnArgPheLeuPheArgLysGlyValSerSerAspTrpValAsnTyr   332
          |||::|:                          |||:
Db   841  AATGGAATAGAGACTAAAACTTTCTTTAGAAAAGAGAATGGAGGATGGAGAGATACT   900
Qy   333  LeuSerProSerGlyValGluArgLeuSerAlaLeuValAspGlyLeuGlySer   352
          |||                             :::
Db   901  TTGAGTGAGTCATTGGCAGGAGAAATTGAATGAAGAGAAGTTAAAGGTTCT   960
Qy   353  GlyLeuThrPhe   356
          |||||||||
Db   961  GGTCTTAAATTT   972
```

```
RESULT 7
US-10-424-599-82242
Sequence 82242, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
  APPLICANT: La Rosa Thomas J
  APPLICANT: Kovalic David K
  APPLICANT: Zhou Yihua
  APPLICANT: Cao Yongwei
  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
  TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
  FILE REFERENCE: 38-21(53223)B
  CURRENT APPLICATION NUMBER: US/10/424,599
  CURRENT FILING DATE: 2003-04-28
  NUMBER OF SEQ ID NOS: 285684
  SEQ ID NO 82242
  LENGTH: 787
  TYPE: DNA
  ORGANISM: Glycine max
  FEATURE:
  OTHER INFORMATION: Clone ID: PAT_MRT3847_45281C.1
US-10-424-599-82242

Alignment Scores:
Pred. No.:
Score:                    725.50        Length:         787
Percent Similarity:       72.98%        Matches:        131
Best Local Similarity:    52.82%        Conservative:   50
Query Match:              38.22%        Mismatches:     62
                                        Indels:         5
                                        Gaps:           2

US-10-019-931-3 (1-359) x US-10-424-599-82242 (1-787)
```

```
Qy    35  LysAlaGluGluAspGluGlyLeuSerCysGluPheGlnGlyGluMetLeuAspSerLeuPro   54
          :::|||||:               |||
Db    49  GAAGCAAGTGGAAGAAAAACAAATTAACCAAGAGTTGTAAGGAGTTAATTCATCTCTTCCT   108

Qy    55  LysGluArgGlyTrpArgThrArgTyrLeuPheThrGlyPheThrTrpCysGlnAla   74
          |||                             :::
Db   109  AGAGAGAAGAGGGTTGGATAACACGTATCTCTATTTATTCCAAGGCTTTGGTACTCATCA   168

Qy    75  LysGluIleGlnIleAlaIleMetSerPheGlyHisPheGlnSerLeuGluAsnAspVal   94
          |||||||                 ::|||
Db   169  ACTGAAAATCCAAGCCATAAAACTTTTCAAAACCAATTCCAAGCCTAAGGACAATGATGTT   228

Qy    95  ValLeuAlaThrIleProLysSerGlyThrThrThrProLeuuLysAlaLeuThrPheThrIle   114
          |||                             :||
Db   229  GTTAATTGCTACTGTTCCAAAATCGGCGCCACCACTTGGTTGAAAAGCCCTTACCTTTGCTATT   288

Qy   115  LeuAsnArgHisArgPheAspProValAlaSerSerThrAsnHisProLeuPheThrSer   134
          |||                             ---
Db   289  CTCCATCGCCAGTATTT-----CCTTCCTTAGAGAACCACCCATTACTTATTTC   339

Qy   135  AsnProHisAspLeuuValProPheGluTyrLysLeuuTyr----AlaAsnGlyAsp   152
          |||||||              ::|||
Db   340  AATCCCCATGAACTTGTGCCTCCAATTTGAATTTGTCATTTATGATGAAAATTAATGCCAA   399

Qy   153  ValProAspLeuSerGlyLeuAlaSerProArgThrPheAlaThrHisLeuProPheGly   172
          |||||||              :::
Db   400  ACTCATGACCTATCCAAAAATGCCCGAGCCCAAGAAATTTTGGGACTCATGTCCATTCACT   459
```

```
RESULT 8
US-10-437-963-55857
Sequence 55857, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
  APPLICANT: La Rosa, Thomas J.
  APPLICANT: Kovalic, David K.
  APPLICANT: Zhou, Yihua
  APPLICANT: Cao, Yongwei
  APPLICANT: Wu, Wei
  APPLICANT: Boukharov, Andrey A.
  APPLICANT: Barbazuk, Brad
  APPLICANT: Li, Ping
  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
  TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
  FILE REFERENCE: 38-21(53221)B
  CURRENT APPLICATION NUMBER: US/10/437,963
  CURRENT FILING DATE: 2003-05-14
  NUMBER OF SEQ ID NOS: 204966
  SEQ ID NO 55857
  LENGTH: 1041
  TYPE: DNA
  ORGANISM: Oryza sativa
  FEATURE:
  OTHER INFORMATION: Clone ID: PAT_MRT4530_57824C.1
US-10-437-963-55857

Alignment Scores:
Pred. No.:
Score:                    674.00        Length:         1041
Percent Similarity:       57.10%        Matches:        139
Best Local Similarity:    38.72%        Conservative:   66
Query Match:              35.51%        Mismatches:     130
                                        Indels:         24
                                        Gaps:           6

US-10-019-931-3 (1-359) x US-10-437-963-55857 (1-1041)
```

```
Qy     1  MetAlaThrSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHis   20
          |||:||                          :||
Db     1  ATGTCTTCCTCCTCCTCCGTGCAAAACCTCTCCC   42

Qy    21  LysLeuGluLeuLysGluGluGlyLysThrArgAspValProLysAlaGluGluAspGlu   40
          ::::|||              |||
Db    43  GAAGTCGAC-----------------------------------TCGCCCTCGCAT   66

Qy    41  GlyLeuSerCysGluPheGlnGluMetLeuAspSerLeuProLysSerGluArgGlyTrpArg   60
          |||                             :::
```

```
Qy   400  ACTCATGACCTATCCAAAAATGCCCGAGCCCAAGAAATTTTGGGACTCATGTCCATTCACT   459

Qy   173  SerLeuuLysGluThrIleThrIleGluLysGlyProGlyValLysValValThrLeuuCysArgAsnPro   192
          :::|::              |||:|:
Db   460  TCATTCGCCAAATCAATTAAAGAGACTGACTGTAGAGTCATTTATATTGTAGGAATCTA   519

Qy   193  PheAspThrPheIleSerSerTrpHisTyrThrAsnAsnIleLysSerGluSerValSer   212
          |||||||              :::
Db   520  TTTGACACTTTCATTTCTTCTTGGCATTACACAAATAATATTAAATCAGAAAGTGTTGCT   579

Qy   213  ProValLeuGluAspGlnIleAlaPheAspProLeuuTyrCysArgGlyValIleGlyPhe   232
          |||||||              :::
Db   580  ACATTACCTCTAGAGGAAGCTTTTGAAAGATTACTGTGAAGGGATAATTGGCTTTGGTCCA   639

Qy   233  PheTrpGluHisIleMetLeuuGlyTyrTrpArgGluSerLeuLeuuAspSerLeuPhe   252
          ||||||              :::
Db   640  TCTTGGAACCATATACTAGATTATTGGAGAGAGCATTGCTAGGCCGAAGAAAGTTTG   699

Qy   253  PheLeuArgTyrGluAspLeuuLysAspAspIleGluThrArgLeuuLysArgLeuuAlaThr   272
          |||||||              |||:
Db   700  TTCTTGAAGTATGAGGACCTTAAAGAAAATGTTCATTTTAATGTGAAGAAAATAGCTGAG   759

Qy   273  PheLeuGluLeuuProPheThrGlu   280
          |||||||
Db   760  TTCTTGGGTTGTCCTTTCACTAAA   783
```

; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO:1749
; LENGTH: 1440
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C40_3
US-10-767-795-1749

US-10-019-931-3 (1-359) x US-10-767-795-1749 (1-1440)

Alignment Scores:
Pred. No.:
Score: 2.12e-71
Percent Similarity: 672.00   Length: 1440
Best Local Similarity: 58.31%   Matches: 136
Query Match: 58.31%   Conservative: 71
DB: 38.31%   Mismatches: 122
                          35.41%   Indels: 26
                          8   Gaps: 8

Qy   8 SerIleProMetAlaIleProSerPheSerMetCysHisHisLysLeuGlnLeuLysGlu 27
              :::||||||||||||::
Db   155 AATATAGTATTTGCACTCATTAGCCGTGTCCCTC------ATGGAATCCCATTCTTGAG 205

Qy   28 GlyLysThrArgAspValPheGlyAspValAlaGluGluGlyGlyLeuSerCysGlnPheGln 47
              ||||||
Db   206 AAACAAAATGGAGATGTG-----------------CTTCAAAAATCTTTCAAA 241

Qy   48 GluMetLeuAspSerLeuProLysGluArgGlyTyrTrpArg---ThrArgTyrLeuTyrLeu 66
              |||::::
Db   242 GAGATGATTTCACTCTCCCTAAAAAGAGTTGTGGGGTTTCCGAAGATCAATATCAA 301

Qy   67 PheGlnGlyPheTrpCysGlnAlaLysGlyIleGlnAlaIleIleMetSerPheGlnLeuHis 86
              |||||:::
Db   302 TATCAAAGTTTTGGTTCAGCCCGCCGTTTCTTCACAGGGCATTGTCGGCCTCAACAACAA 361

Qy   87 PheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLeuLysSerGlyThrThrTrp 106
              |||:::|:
Db   362 TTCCAAGCTCAGCCAACTGATATCATCCTTTGTTAGCTCTCCAAGAACAGGACAGCCTGG 421

Qy   107 LeuLysSerAlaLeuThrPhePheThrIleLeuAsnArgHisArgPheAspProValAlaSerSer 126
              ||:::||::
Db   422 TTAAAACCCTCACTTTCGGCCACTTATACAGAGAAACTTCATAC----------AATGAT 469

Qy   127 ThrAsnHisProLeuPheThrSerAsnProHisAspLeuValProPhePheGluTyrLys 146
              ||||||::
Db   470 TCCACCACCCCTTTACTTTCCAAGATGCCCTCATGAGTTGTCCCTTACATGGAGTTTGAT 529

Qy   147 LeuTyrAlaAsnGlyAspValProAlaAspLeuSerGlyLeuValAlaSerProArgThrPheAla 166
              ||:::
Db   530 ------CATGCCCAGATTCCCACTAACTTGGATGGAGCAGCATCTGGAATTCCT---CTTTTAGCC 577

Qy   167 ThrHisLeuProPheGlySerLeuLysGluThrIleGlyLeuProGlyValLysValVal 186
              ||:::||::
Db   578 ACTCACTCCTTCTTATTCCTTCTTACCCAGATCATAATTGGGTTGTTGTAAACTTATT 637

Qy   187 TyrLeuCysArgAsnProPheAspThrPheIleSerSerPheHisTyr---ThrAsnAsn 205
              ||:::||
Db   638 TACATTTGCAGGAGCCCCAAAGATACATTTGTTCATTGTATCATTCAGCGAC 697

Qy   206 IleLysSerGluSerValSerProValLeuLeuAspGlnIleAlaPheAspLeuTyrCysArg 225
              ||||||::
Db   698 TGCAAATCCCAAAATGCTCAACCGTTCAACTTGATGAAGCGGTTCGAATTGTTTTATGAA 757

Qy   226 GlyValIleGlyPheGlyProThrTrpGluHisMetLeuGlyTyrTrpArgGluSerLeu 245
              ||:::||
Db   758 GGTGTAAGCCCGTTGGCCTTATTGGACCAATGTTTGGGGTACTGGAAAGCAAGTTTC 817

Qy   246 LysArgProGlnLysValPhePheLeuArgTyrGluAspLeuLysAspAspIleGluThr 265
              ||:::
Db   818 GAACAATCCGGACAAGTTTATGTGTTCTTGAAAAATGAAGAAGATACTGTTTTG 877

RESULT 9
US-10-767-795-1749
Sequence 1749, Application US/10767795
Publication No. US20040181830A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Cao, Yongwei
APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(53534)B

Qy   67 GAACTCTACAAGCAGTTCACCGACTCGGTGTCTCCCTGCCGTGAGCTCAGAACCCATGCCC 126
              |||||::
Db   61 ThrArgTyrLeuTyrLeuPhe---GlnGlyPheTrpCysGlnAlaLysGlyIleGlnAla 79

Db   127 TTCCTCCCGCTTACCGCACGACAAAGGCTGGGTACAGTAGCCTCATTGGTCGGT 186

Qy   80 IleMetSerPheGlnLeuHisPheGlnSerLeuGluAsnAspValValLeuAlaThrIle 99
              :::||
Db   187 GCCGATGGTCGCCGACGGCGGCGGTTCGCCGCGGCGCCCCTCCGACATCATCGTCGCCACCTTG 246

Qy   100 ProLysSerGlyThrThrTrpLeuLysSerAlaLeuThrPhePheThrIleLeuAsnArgHisArg 119
              ::::|||
Db   247 CCCAAGTCCGGCACGACGTGGATCAAGGGCGGCTCCTCTACCGCCCACGGGTGCACCCG---AGG 303

Qy   120 PheAspProValAlaSerSerThrAsnHisProLeuPheThrSerAsnProHisAspLeu 139
              :::::||
Db   304 GAGCACCCGCGGCCGACCGCCGCGGCGCCGCCGCCCTTCAACTCCCTGCCCCCCACGAGTGC 363

Qy   140 ValProPhePheGluTyrLysLeuTyrAlaAsnGlyAspValProAlaAspLeuSerGlyLeu 159
              :::
Db   364 GTCAACTTCCTCGGATCACCACAGCTCTTACCACCAACAACAGAGTCCCCGGACCTGGGCAGGCTG 423

Qy   160 AlaSerProArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGlu 179
              ||||||
Db   424 CCCGGACCCGGAGGCGTGTTCCGGACCAGTCCGTTCACGTCGCTGCGGAGCGGCCCGCG 483

Qy   180 LysProGlyValLysValValTyrLeuCysArgAsnProPheAspThrPheIleSerSer 199
              :::::|::
Db   484 GGGTCGGGGTGCAAGGTCGTGTACGTGTCGCGCAACCCGGACGCCACCTTCATCTCCAGT 543

Qy   200 TrpHisTyrThrAsnAsnIleLysSer---GluSerValSerProValLeuLeuAspGln 218
              ::::
Db   544 TGGCACTTCCCACAACGGTTCAGGGCCCGTGAGGGGCACGGCCAGGCC 603

Qy   219 AlaPheAspLeuTyrCysArgGlyValIleGlyPheGlyProThrTrpGluHisMetLeu 238
              ||||||
Db   604 ATCGCCGAGCTGTTCTGCCTCGGCGTGCCCGGGTGCGGGTACTGGGAACCAGTCCTC 663

Qy   239 GlyTyrTrpArgGluSerLeuLysArgProGlnLysValPhePheLeuArgTyrGluAsp 258
              ||||||:::
Db   664 GGCTACTGGGCGCGGGAGCGCGTGGGCCGCCCCGAGCAGGTCCTCTTCTTCAGGTACGAGGAG 723

Qy   259 LeuLysAspAspIleGluThrArgAsnLeuLysAlaThrPheGluGluLeuProPhe 278
              :::|
Db   724 ATGAAGCTCGACGCGCGCGCACGTCCCGGAGGCAGGTCGTCGGCCTGCCGTTC 783

Qy   279 ThrGluGluGluGlyGlyLeuValValLysAlaAlaAlaIleAlaGluLeuCysSerPheGlu 298
              ::
Db   784 AGCGCCGAGGAGGAGGGAGGGAAGGTGGTGGGACGGCCATCGCAGCCTGTGCTCGTTCGAT 843

Qy   299 AsnLeuLysGluValAlaAsnLysSerAsnLysSer------IleLysAsnPhe 315
              :::::
Db   844 CACAGCATGCCCTGGAGCCACCAAGAGCGGGTCGTCGGCACGGGG 903

Qy   316 GluAsnArgPheLeuArgLysGlyLysGlyValSerAspTrpValAsnTyrLeuSerPro 335
              :::
Db   904 GCGAATAGCTCATTCTTCCGTCGCGGACTGGGACCGAACCATCTTTCGCCG 963

Qy   336 SerGlnValGluArgLeuSerArgAlaLeuValAlaAspApLysLeuGlyLysGlyLeu 354
              :::|::
Db   964 GAGATTGCACGCGCGGATCACGCCATAACCGAGTTCAATGGTTCCGGGTCTC 1020

Qy   266   AsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGluGluGluArgLys   285
           ||||||:|:|||:|| ||||:|||||:||:||||
Db   878   TATCTTAAGAAATTGGCAGAGTTTATGGGGTTGTCCTTTCTCAACAGAGGACAGCAAAA   937

Qy   286   GlyValValLysAlaIleAlaGluLeuCysSerPheGluAsnLeuLysLysLeuGluVal   305
           :|:|| ||||| ||:||||||||||||||
Db   938   GGGGTACCTGAAAAAGATTATAAAAGATGTGTAGTTTTGAGAATCTAAGCAACTTGAAGTG   997

Qy   306   AsnLysSer------------------AsnLysSerIleLysBAsnPheGluAsnArgPheLeu   320
           ||||
Db   998   AATAAAAGTGGGATAACATCGTTGGAAGGGCAACGGAATTTGGAGATTCAAAACAAGATTAT   1057

Qy   321   PheArgLysGlyGluValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArg   340
           |||||::|||||| ||||:::||||
Db   1058   TTTCGGAAAGGGAGGATTGGAAGAATTATTTGACCCCTAAAATGCCACACAG   1117

Qy   341   LeuSerAlaLeuValAspAspLysLeuGlyGlySerGlyLeuThr   355
           ::: |||||
Db   1118   TTAGACAAAATAACTCAGCAAAAAATTAAGTAGCTCAGGTTTAAGT   1162

RESULT 10

Alignment Scores:

Pred. No.:	1.13e-67	Length:	271990
Score:	671.50	Matches:	139
Percent Similarity:	61.49%	Conservative:	75
Best Local Similarity:	39.94%	Mismatches:	122
Query Match:	35.38%	Indels:	13
DB:	6	Gaps:	5

US-10-019-931-3 (1-359) x US-10-195-144-87 (1-271990)

Qy   15   SerPheSerMetCysHisLysLeuGluLeuLysGluGlyLysThr--------------   30
           |||
Db   244933   AACATATCCTTTGTAACGAGTAACAAGAGAACAACAAATGGATCCG   244874

Qy   31   ArgAspValProLysAlaGluGluAspGlyLeuSerCysGluPheGlnGluMetLeu   50
           ||||
Db   244873   AAGGAGCTTCCGTCGAACCTTAGAGACGACAAGAGAAGTGAAGAAAACCAGAAGGTGATC   244814

Qy   51   AspSerLeuProLysGlyLeuArgGlyTyrTrpArgThrArgTyrLeuTyrLeuPheGlnGlyPhe   70

Db   244813   TCTTCGCTTCCTTCACACACGACAGATTACCAAGGCATAAGATTGTAAAATATCAAGGATGT   244754
           :|||||||

Qy   71   TrpCysGlnAlaLysGluIleGlnIleAlaIleMetSerPheGlnLysHisPheGlnSerLeu   90
           :::|| ||||::|:
Db   244753   TGGTATTATTACAACACACCTCCAAGGTGTCTC-AATTTCCAGAGAGGGTTTTCAAACCGCAA   244695

Qy   91   GluAsnAspValValAlaThrIleProLysSerGlyThrThrThrTrpLeuValSerAlaLeu   110
           :::|| ||||:|:::|
Db   244694   GACACTGCATGTAATCATTGCTTCGTACCCCAAAATCAGGCACTTACTGGCTCAAGGGCCTC   244635

Qy   111   ThrPheThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThrAsnHisPro   130
           |||||:::
Db   244634   ACAGTGCGCTCTGCTTGAGAGGATCAAAG----------AACCACTCTTCTGATCACCT   244587

Qy   131   LeuPheThrSerProHisAspLeuValProPhePheGluTyrTyrLysLeuTyrAlaAsn   150
           |||:|:|||||||
Db   244586   CTCCTATATCATAATCCTCATGGCATTATACCCATTCTTGGAGATCGGATGTGTACCACGAA   244527

Qy   151   GlyAspValProLeuSerGlyLeu---AlaSerProArgThrPheAlaThrHisLeu   169
           :::|:||:|||
Db   244526   AGCTCAAGTCTTAACCTAGCCAAGTTCTCAGCACCTCCGAGGCTGTTCTCGACTCACATG   244467

Qy   170   ProPheGlySerLeuHisGluThrIleGlyLysProGlyValLysValValTyrLeuCys   189
           :::|:|
Db   244466   CCACTGCACACGATCCACGAAGCACTCAAGCACTCTCCTTGCAAGGTGTGTTACGGTGTGC   244407

Qy   190   ArgAsnProPheAspThrPheIleSerSerTrpHisTyrThrAsnAsnIleLysSerGlu   209
           |||:::|
Db   244406   AGGAACGTGAAGGACAGCAGGTTGGTTTAGAGCTGTTTGGTTTACAGCTGTGCTATATAAAATC   244347

Qy   210   SerValSerProValLeuLeuAspGlnAlaPheAspLeuTyrCysArgGlyValIleGly   229
           :::|:||||||
Db   244346   GAACCAACCAGAGAGTTCTCGAGTCTATGTTTAAGCGAGTTCTGCGGATGGAACCAACTAT   244287

Qy   230   PheGlyProPheTrpGluHisMetLeuGlyTyrTyrArgGluSerLeuLysArgProGlu   249
           |||:::
Db   244286   TTTGGACCTTTTTGGGATCATCTCTCTCTGGTAATCTGAGTTACTGGAGACCCAAAG   244227

Qy   250   LysValPhePheLeuArgTyrGluAspAspAlaIleGluThrAsnLeuLysArg   269
           |||:|:|||
Db   244226   CATGTCCTTTTCATTGAGGTATGGAGGATGAAAGCCGAGCCTTCGTGATCAGATCAAGAGA   244167

Qy   270   LeuAlaThrPheLeuGluLeuProPheThrGluGluGluArgLysGlyValValLys   289
           ::|||||||
Db   244166   CTTGCGGACTTCTTGGGATGTCCTTTTACTAAGCAACAGAGAAGAGATAGTGGATCGTGTGGAC   244107

Qy   290   AlaIleAlaGluLeuCysSerPheGluAsnLeuLysLysLeuGluValAsnLysSerAsn   309
           :::|||:||||:|
Db   244106   GGGGATCTTGGACCTCTCCTCTCGTAATCTGACGAGTTTGGAGGCTAACAAAAACAGGG   244047

Qy   310   LysSerIleLysAsnPheGluAsnArgPheLeuPheArgLysGlyGluValSerAspTrp   329
           :::|||
Db   244046   --ACAATAAACAAATGTGGGAGCACAAGTTTTTTCCGTAAAGGAGAAGTCGGTGACTCG   243990

Qy   330   ValAsnTyrLeuSerProSerGlnValGlu-ArgLeuSerAlaLeuValAspAspLysLeu   349
           |||||:::
Db   243989   AAAAATATTCTTACG---TCTGAAAATGGAGAAAGATAGACATGAATCATCAAAAGAAAAACT   243933

Qy   349   uGlyGlySerGlyLeuThrPhe   356
           :::|||||||
Db   243932   TCAAGGTTCTGGTTTGAGTTTT   243911

RESULT 11

APPLICANT: JIN, HUA
APPLICANT: LAI, FANG MING
APPLICANT: LEFOREST, MARTIN
TITLE OF INVENTION: NUCLEAR FERTILITY RESTORER GENES AND METHODS OF USE IN
TITLE OF INVENTION: PLANTS
FILE REFERENCE: 16313-0210
CURRENT APPLICATION NUMBER: US/10/345,072
CURRENT FILING DATE:
PRIOR APPLICATION NUMBER: US/10/345,072
PRIOR FILING DATE: 2003-01-16
PRIOR APPLICATION NUMBER: PCT/US02/22217
PRIOR FILING DATE: 2002-07-12
PRIOR APPLICATION NUMBER: 60/305,026
PRIOR FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: 60/305,363
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: 60/305,363
PRIOR FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: 60/308,736
PRIOR FILING DATE: 2001-07-30
NUMBER OF SEQ ID NOS: 179
SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 87
LENGTH: 271990
TYPE: DNA
ORGANISM: Raphanus sativum
FEATURE:
NAME/KEY: modified_base
LOCATION: (144241)..(144300)
OTHER INFORMATION: a, t, c, g, other or unknown

US-10-345-072-87

Alignment Scores:
Pred. No.:
Score: 671.50   Length: 271990
Percent Similarity: 61.49%   Matches: 139
Best Local Similarity: 39.94% Conservative: 75
Query Match: 35.38%   Mismatches: 122
                             Indels: 13
                             Gaps: 5

US-10-019-931-3 (1-359) x US-10-345-072-87 (1-271990)

```
Qy     15 SerPheSerMetCysHisHisLeuGluLeuLeuLysGluGlyLysLysThr----------- 30
          |||  :::       ::: ||| :::||| ||||             :::|||
Db 244933 AACATATCCTTTTGTAACGAGTAACAGAGTCAAGAACCAAGAAACACAAATGGATCCG 244874

Qy     31 ArgAspValProLysAlaGlyAlaAspGluGlyLeuSerCysGlyPheGlnGluMetLeu 50
          |||:|||       :::  |||   :::|||       :::
Db 244873 AAGGAGCTTCCGTCGAACCTTAGACGACGAAGAAGTAGAAGAAACCAGAAGGTGATC 244814

Qy     51 AspSerLeuProLysGluArgGlyTrpArgThrArgTyrLeuTyrLeuPheGlnGlyPhe 70
          |||||||||       :::   :::|||
Db 244813 TCTTCGCTTCCTTCACACAGATTACCAAGGGCATAAGATTTGTAAATATCAAGGATGT 244754

Qy     71 TrpCysGlnAlaLysGluIleGlnIleIleMetSerPheGlnLysHisGlnSerLeu 90
          :::||||||       :::   |||:::
Db 244753 TGGTATTATTACAACACCTCCAAGGTGTCTC-AATTCCAGAGAGGTTTTCACCGGCAA 244695

Qy     91 GluAsnAspValValLeuAlaThrIleProLysSerGlyThrThrTrpLeuLysGlnAlaLeu 110
          :::|||||||:::
Db 244694 GACACTGAAGTAAACATTGCTTTCGTACCCCAAATCAGACGACTACTTGGCTCAAGGCCCTC 244635

Qy    111 ThrPheThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThrAsnHisPro 130
          :::|||       :::       ::: |||
Db 244634 ACAGTCGCTTCGCTTGCTTGAGAGATCAAAG--------AACCACTCTTCGATCATCCT 244587

Qy    131 LeuPheThrSerAspProHisAspLeuValProPheGluTyrLysLeuTyrAlaAsn 150
          |||||||    :::    :::||| :::
Db 244586 CTCCTATATCATAATCCCATGGCATTATACCAATTCTTGGAGATCGATGGTACCACGAA 244527

Qy    151 GlyAspValProAspLeuSerGlyLeu---AlaSerProArgThrPheAlaThrHisLeu 169
          :::       :::|||   :::|||
Db 244526 AGCTCAAGTCCTACCTAGCCAAGTTCTCAGCACCTCCGAGGCGTCGTTCTCGACTCACATG 244467

Qy    170 ProPheGlySerLeuLysGluThrIleGlyLysProGlyValLeuSerValValTyrLeuCys 189
          |||       :::|||:::||| :::||| :::
Db 244466 CCACTGCACACGATCCACGAAGCACTCAAGCACTCTCCTTGCCAAGAATGTTGTACGTGTGC 244407

Qy    190 ArgAsnProPheAspThrPheIleSerSerTrpHisTyrThrAsnAsnIleLysSerGlu 209
          |||||||       :::|||       :::
Db 244406 AGGAACGTGAAGGACACGTTCATCTCGTCGTGGCATTACACGAACAACATAAAATC 244347

Qy    210 SerValSerProValLeuLeuAspGlnAlaPheAspAspLeuTyrCysArgGlyValIleGly 229
          :::||| :::       ||| :::|||
Db 244346 GAACCAACCAGAAGAGTTCTCCGAGCTATGTTAACGAGTTTCTGCGATGGAACCAACTAT 244287

Qy    230 PheGlyProPheTrpGluHisMetLeuGlyTyrTyrArgGluSerLeuLysArgProGlu 249
          |||       ::: |||   :::|||
Db 244286 TTTGGACCTTTTTTGGGATCATCTCTTGAGTTACTGCGATGGGAAGACCCAAAG 244227

Qy    250 LysValPheGluLeuArgTyrGluAspLeuLysAspAspIleGluThrArgAsnLeuIleArg 269
          |||       :::|||   :::||| :::
Db 244226 CATGTCCTTTTCATGAGGTATGAGGACCTGAAAAGCCGAGCCTCGTGATCAGATCAAGAGA 244167

Qy    270 LeuAlaThrPheLeuGluLeuProPheThrGluGluGluGluGluLysGlyGlyValValLys 289
          |||       :::|||:::||| :::
Db 244166 CTTGCGGACTTCTTGGGATGTCCCTTTTACTCAAGAACAAGAAGAAGATCTGTGGAC 244107

Qy    290 AlaIleGluAlaGluLeuCysSerPheGluAsnLeuLysGluLeuValAlaAsnLysSerAsn 309
          |||       :::   |||   :::|||
Db 244106 GGGACTGGACCTCTGCTCTCGTCCTAATCTGAGCAGTTTGGAGCGTAACAAAAACGGG 244047

Qy    310 LysSerIleLeuLysAsnPheGluAsnArgPheLeuProPheArgGlyLysGlyValSerAspTrp 329
          :::||| :::       :::|||:::
Db 244046 ---ACAATAAACAATGTGGAGCACAAGTTTTTTCCGTATAAGGAGAAGTCGGTGACTCG 243990

Qy    330 ValAsnTyrLeuSerProSerGlnValGlu-ArgLeuSerAlaLeuValAspAspLysIle 349
          |||   :::       :::||| :::
Db 243989 AAAAATTATCTTACG---TCTGAAATGAGAAAAGAGACATGATCATCAAAGAAAAAACT 243933

Qy    349 uGlyGlySerGlyLeuThrPhe 356
          :::|||       :::|||
Db 243932 TCAAGGTTCTGGTTTGAGTTTT 243911
```

RESULT 12
US-10-767-795-1750
; Sequence 1750, Application No. US20040018830A1
; Publication No. US20040018830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 1750
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C40_2
US-10-767-795-1750

Alignment Scores:
Pred. No.:
Score: 666.50   Length: 1251
Percent Similarity: 58.19%   Matches: 137
Best Local Similarity: 38.70% Conservative: 69
Query Match: 35.12%   Mismatches: 120
                             Indels: 28
                             Gaps: 8

US-10-019-931-3 (1-359) x US-10-767-795-1750 (1-1251)

```
Qy 16 PheSerMetCys---His-LysLeuGluLeuLeuLysGluGlyLysThrArgAspValPr 34
       :::||| |||       :::|||
Db 40 TATAGTATTTCGCACTCATTAACCCTGTCTCTCATTGAATCCCACTTTGATTCCCATGTGA 99
```

RESULT 13
US-10-739-930-1663
; Sequence 1663, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
;

Alignment Scores:
Pred. No.:   1.05e-70   Length: 1358
Score: 666.00   Matches: 138
Percent Similarity: 58.33%   Conservative: 58
Best Local Similarity: 41.07%   Mismatches: 124
Query Match: 35.09%   Indels: 16
DB: 8   Gaps: 8

US-10-019-931-3 (1-359) x US-10-739-930-1663 (1-1358)

Qy   34   oLysAlaGluGluAspGluGlyLeuSerCysGluPheGluMetLeuAspSerLeuPr   54
Db   100   GCAACTAACGAAGATGAG---TTTCAGAAATCTTCAAAGAGATGATTTCTACTCTCCC   156

Qy   54   oLysGluArgGlyTrp---ArgThrArgTyrLeuTyrLeuPheGlnGlyTrpCysG1   73
Db   157   TAAAGGAACAGCTGGGGGTCTTCCCGATCATCTATATCAGTATCAAGATATTTGGTTCAC   216

Qy   73   nAlaLysGlurleGlnAlalleMetcSerPheGlnLysHisPheGlnSerLeuGluAsnAs   93
Db   217   CTCACCTTTCCTACAAGGAGTGCCTGTTAGCTCAACAACAGTTTGAGGCTCAACCCACTGA   276

Qy   93   pValalValLeuAlaThrIleProLysGlyThrThrTrpLeuTyrAlaLeuThrPheTh   113
Db   277   TATCATACTTTCTAGTGCCCCTAGAACCGGCACAGCCTGGTTAAAAATCTCACTTTCGC   336

Qy   113   rIleLeuAsnArgHisArgPheAspProValAlaSerSerThrArgAsnHisProLeuPheTh   133
Db   337   TACTATTACAAGAACTTCATACGAT------GATTCTACCACCCCTTTGCTTTT   384

Qy   133   rSerAsnProHisAspLeuValProPhePheGlyTyrLysLeuTyrAlaAsnGlyAspVa   153
Db   385   CAGGATGCCCATCAGAGTTGTGCCTTTCATCGGAGCTTT------   421

Qy   153   1ProAspLeuSerGlyLeuAlaSerProArg---------ThrPheAlaThrHi   168
Db   422   ------GATCATGCTCATTTTCCGCTAATCGACATCTTGGAATCCTTCTTTGGGCCACTCA   477

Qy   168   sLeuProPheGlySerLeuLysGluThrLeuGlyLysProGlyValLysValValTyrIle   188
Db   478   TGCTCCTTATTCTTTCTTACCTACATCAATAATTGATTCTCGTTGTTGTAAAATTATTTACAT   537

Qy   188   uCysArgAsnProPheAspThrPheIleSerSerThrPHisTyr---ThrAsnAsnIleTy   207
Db   538   TTCGCAGGGACGACCCAAGGATACATTTGTTCTTGAAATTATGAACATCTGTACCAGGTATCCGAA   597

Qy   207   sSerGluSerValSerProValLeuLeuAspGlnAlaAlaPheAsnLeuTyrCysArgGlyVa   227
Db   598   ATCCCAAAACACTCAACCCATTCAACTTGATGAAGGCTCGAGTTATTTATGAAGGGT   657

Qy   227   lIleGlyPheGlyProPheTrpGluHisMetLeuGlyTyrArgTrpArgSerLeuLysAr   247
Db   658   AAGTTGGTATGGGCCTTTATTGGGACCCATAGTTCTCGGGGTACTGGAAAGCAAGACTTGGAACA   717

Qy   247   gProGluLysValPhePheLeuArgThrGlnAspLeuLysAspAspArgIleGluThrAsnLe   267
Db   718   TCCAGACAAGTTAAGTTCGTTCTTGAAATAGAAGAACTGATGTAGAAACTACTGTTTGTATCT   777

Qy   267   uLysArgLeuThrPheLeuGluLeuProThrGluGluGluArgLysGlyVa   287
Db   778   TAAGAAAAATAGCAGAATTTATTGGGTTATCCTTTCTCGTCAGAGGAACAACAAGGGGT   837

Qy   287   lValLysAlaIleAlaGluLeuCysSerPheGluAsnLeuLysLysLysLeuValAsnLy   307
Db   838   GCCTGAAAAACATTGTACAGCTGTGCAGAATTTAAGTGGCTTGGAAGTAAATAA   897

Qy   307   sSerAsnLysSerIleLysAsn---------PheGluAsnArgPheLeuPheAr   322
Db   898   AACTGGGAAACATCCTGACGCGGCCAAGGAGATTGGGATGAAAAAATAACATTTTCTTCCG   957

Qy   322   gLysGlyGlyValSerAspTrpValAsnTyrLeuSerProSerGlnIleArgLeuSe   342
Db   958   GAAAGGGAAGGTTGGAGACTGGAAGAATTATTTGACAACTGAAATGGCTCGACGTTTAGA   1017

Qy   342   rAlaLeuValAspAspLysLeuCysGlySerGlyLeuThr   355
Db   1018   CCAACGAAACTGCAAAAGTTTCGCGGGTTCAGGTTTAAGT   1057

Qy   34   ProLysAlaGluGluAsp---------GluGly------------LeuSerCysGluPheGln   47
Db   179   CCCGGAGCAAGAAGACCGATGCGGAAACGAACCCTGAGCTCTACCAGCATTTCGCC   238

Qy   48   GluMetLeuAspSerLeuProLysGluArgGlyTrpArgThrArgTyrLeuTyrLeuPhe   67
Db   239   AACTCGGTTCCTCCCTCCCTGCGCCGAGCTCCAAAGGCTTGGTCCAACAACAAATTCTACCGCAC   298

Qy   68   ---GlnGlyPheTrpCysGlnAlaLysGluIleGlnIleGlnAlaIleMetSerPheGlnLysHis   86
Db   299   GACCAAGGCTCGGACACTGCACGCTGGTGGGGCATAGTGGGTCGCCGGACGGGTGC   358

Qy   87   PheGlnSerLeuGluAsnAspValValLeuAlaThrLeuProLysGlyThrThrTrp   106
Db   359   TTCAGCGGCGGCCGCCCCTCGGAGTCATCGTCGCCGACCACCCGCCAAGTCCGGACCACGTGG   418

Qy   107   LeuLysAlaLeuThrPheThrIleLeuAsnArgHisArgPheAspProPhePheGlu---Ala   124
Db   419   ATGAAGGTCGCCTCCGTGACGCCAACCGTCACCCG--------AGAGACGTCCCGTGCTGCGGGC   475

Qy   125   SerSerThrArgAsnHisProLeuPheThrSerAsnProHisAspLeuValProPhePheGlu   144
Db   476   GCCGCGCCGACCAGCATCCTCTCCAACAGCTGGGCGCCCCCCACGGAACTCATCAAGTTCTTCGAG   535

Qy   145   TyrLysLeuTyrAlaAsnGlyAspValProLeuSerGlyLeuAlaSerProArgThr   164
Db   536   TACCAGCTCTACACGGGGACAAGGGTCCCGGACACCTCCCGTGAGGCCCCCCGGCTC   595

Qy   165   PheAlaThrHisLeuProPheGlySerLeu---LysGluThrLeuGlyLysProGlyVal   183
Db   596   TTCGCGACGCACGTCCCCTTTGGTGTCCCACGGCTGC   655

Qy   184   LysValValTyrLeuCysArgAsnProPheAspThrPheIleSerSerSerThrHisTyrThr   203
Db   656   AAGATCGTCTACGGTGTGCGCAACGACACCTTCCGTGTCCGGGGACTTCGTC   715

Qy   204   AsnAsnIle------LysSerGluSerValSerProValLeuLeuAspGlnAlaAlaPheAsp   221
Db   716   AACAAATACTTCAGGGTGCAAGGAGTGGCCTCGACCTCTCCGTGAGGCCCCCCGGA   775

Qy   222   LeuTyrCysArgGlyValIleGlyPheGlyProPheTrpGluHisMetLeuGlyTyrTrp   241
Db   776   CTGTTCTCGGCGCGGCCCTACATCGGGTTCGGGCCGTTCTGGGAGCACATGCTCGGCTACTGG   835

Qy   242   ArgGluSerLeuLysArgProGluLysValPhePheLeuArgThrGluAspLeuLysAsp   261
Db   836   CGCGGGAGCCTGAAGAGGCCGGAGAAGGTGTTCTTCCTGCGCACGGAGGACCTGAAGGAC   895

Qy 262 AspIleGluThrAsnLeuLysArgLeuAlaThrPheLeuGluProPheThrGluGlu 281
        |||||||| :|::|||
Db 896 GACCCCGCGCCGCGACGTCCCAGGCGTTGGCGAGTTCGTCGTCCCAATTCAGCTCCGAG 955

Qy 282 GlyGluArgLysGlyValValLysAlaIleAlaGluLeuCysSerPheGluAsnLeuLys 301
        ||||| :|||
Db 956 GAGGAGGACGGAGCGGCGCGGTGGTGGTGTCGTGTCGCGTGTGTGGGCCATGAGG 1015

Qy 302 LysLeuGluValAlaAsnLysSerAsnLysSer------------IleLysAsnPheGluAsnArg 318
        :::|: ||| |
Db 1016 GGCCATGGCCGGACCAAGGGCGGACGAAGTTCCAGACGTTCCCAACAGC 1075

Qy 319 PheLeuPheArgLysGlyValValSerAspTrpValAlaAsnTyrLeuSerProSerGlnVal 338
        || |:: ::
Db 1076 TCCTTCTTCCGGCAGGGCGTGGTGGGGTGGGGTGAACCACACTGTCACCGGAGATGGCG 1135

Qy 339 GluArgLeuSerAlaLeuValAlaAspAlysProLeuGlyGlySerGlyLeu 354
        ||||::| ||
Db 1136 CGCGCGCCATTGACGCGGATCACGCAGACAAGTTCAGGGGTTCGAGTCTC 1183

RESULT 14

US-10-437-963-55524

; Sequence 55524, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 55524
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_57523C.1

US-10-437-963-55524

US-10-019-931-3 (1-359) x US-10-437-963-55524 (1-1077)

Alignment Scores:

Pred. No.:

Score: 4.57e-70
Percent Similarity: 58.81%
Best Local Similarity: 41.49%
Query Match: 34.75%
DB:

	Length:	1077
	Matches:	139
	Conservative:	58
	Mismatches:	125
	Indels:	13
	Gaps:	5

Qy 32 AspValProLysAlaGluGluAspGluGly------------LeuSerCysGluPheGlnGlu 48
     || |||
Db 55 GACGACATGAGCGCGCCGGCCAAGGCGGAAGGCTGGTCGTTCAACCCCTGCCAAGTACGAGGCG 114

Qy 49 MetLeuAspSerLeuProLysGluArgGlyTrpArgThrArgTyrLeuTyrLeuPheGln 68
     ||||||||
Db 115 CTCGGCGTCCTCCCTACCGTTCTTGCCATGGTCTCCGGAGCGGCCCCGTACCGCAGGTACGGC 174

Qy 69 GlyPheTrpCysGlnAlaLysGlnIleGlnIleAlaIleMetSerProHisGlnLysPheGln 88
     ||| ||||
Db 175 GGATTCTGGTGCCCGGCCAGCACCTGGCCGCCGGGACGTTGGTG 234

Qy 89 SerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThrTrpLeuTyr 108
     ::: ||
Db 235 GCCCGCCCCAGCGAGGTCATCCTCGCCACCATACCCAAGTCCGGCTCGAGGTGGCTCAAG 294

Qy 109 AlaLeuThrPheThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThrAsn 128
     :::|||: ||
Db .295 GCGCTCGCCTTCTCGGTCGTTGCACCGCGGTGCACCGGCCGACCGCCGCGGG------CAG 348

Qy 129 HisProLeuPheThrSerAsnProHisAspLeuValProPhePheGlu------TyrLys 146
     ||| :::
Db 349 CACCCGCTCTTCCACTCCACCCCCGACCTCGTGCCATTCCTCGACTACAAG 408

Qy 147 LeuTyrLeuAlaGlyAspValProAsp------------LeuSerGlyLeuAlaSerProArg 163
     ||| :::
Db 409 ATCAGCGCGGACTCGCCGCCGTCCCCGCCCAGGCCATGCCGTCCGGC 468

Qy 164 ThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGluLysProGlyVal 183
     ||||| ||
Db 469 ATCCTGGCAGTGCACGAAACCGCTCTCGAGCCTCGCCCCGGTCAAGGCGGTTCCGGGTGC 528

Qy 184 LysValValLysTyrLeuCysArgAsnProPheAspThrProPhePheIleSerSerTrpHisTyrThr 203
     ::: |||||:
Db 529 CGCGTCGTTCACCTCTGCCGCGACCCCAACGGACGGGTTCCTCCGCTCCGGCATTACCTC 588

Qy 204 AsnAsnIleLysSerGluSerValSerProValLeuLeuAsnGlnAlaPheAspLeuTyr 223
     :::|||::| ||
Db 589 GACGAAGATCAAAACCAGAGGGAATCCACCATGACGCCGGTTCGACGAGCCTTCGACGTGCTC 648

Qy 224 CysArgArgLysGlyValIleGlyPheGlyGlyProThrTrpGluHisMetLeuGlyTyrTyrArgGlu 243
     ||||||
Db 649 TGCCGACGGCGTCTCGCGGCCTATGTGGGGACCACCGGCGCGCCGAGTACTGGAACGAG 708

Qy 244 SerLeuLysPheArgProGluTyrValPhePheLeuArgTyrGluAspLeuLysAspAspAlaIle 263
     |||::|:||
Db 709 AGCCTGGCGGGGCGGCGGGAGGTCGTTGTTCCTCGGCGGTCCACGAGGCGGACGACGGC 768

Qy 264 GluThrAsnLeuLysArgLeuAlaThrPheLeuGluProPheThrGluGluGlu 283
     |||||||||:|:::||||
Db 769 GTGGGGAGGGTGAGCGCGGCGGCGCGCTCTTCGGCGGCGTTCCCGGTTCACCGGGAGCTT 828

Qy 284 ArgLysGlyValValLysAlaIleAlaGluLeuCysSerPheGluAsnLeuLysSerLeu 303
     ||||:::||
Db 829 GCCGGTGGCGTCCCCGAAACTATCGGTGGAGTCTCGGAGTATGGAAAGGATGAGAAAGTG 888

Qy 304 GluValAlaAsnLysSerAsnLys------------SerIleLysAsnPheGluAsnArgPheLeu 320
     ||| ::: :::
Db 889 GAGGCAAATCGAGACGGCGAGCCACGGGGCGTACCTGGTCGTTCAAGAACTCGGCCTTC 948

Qy 321 PheArgLysGlyValValSerAspTrpValAlaAsnTyrLeuSerProSerGlnValArg 340
     |||||||||||||::|
Db 949 TTCAGGAAGGGTGAGGTCGGAGACTGGAAGGAGCACATGTCGCCGGAAGATGGCGCGGAGG 1008

Qy 341 LeuSerAlaLeuValAlaAspAlaLysProLeuGlyGlySerGlyLeuThr 355
     ||||:::||
Db 1009 CTCGAGGACGTGCGTGGAGGAGAAGCTGCGAGGCTCCCGGATGTCG 1053

RESULT 15

US-10-425-115-167844

; Sequence 167844, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 167844
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_84655C.1
;

US-10-425-115-167844

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

Length: 1579
1.91e-68
648.50
55.84%
38.75%
34.18%
8

Matches: 136
Conservative: 60
Mismatches: 142
Indels: 13
Gaps: 6

US-10-019-931-3 (1-359) x US-10-425-115-167844 (1-1579)

```
Qy   14 ProSerPheSerMetCysHisLysGluLeuLeuLysGluGlyLysThrArg------  31
                                                        :::::
Db  277 CCATCGTTCTCCTGAACCAAAGCCATGGGCGCAAGCTGCTTCCGGCTCCAGCAAGATCGAC 336

Qy   32 ---AspValProLysAlaGluGluAspGlyGlyLeuSerCysGluPheGlnGluMetLeu  50
           :::                      ::::::::
Db  337 ACCAAAGCTCCACTCAGCCGACGAGGCCGCTCAACCCAAGAGCTACGGCGACTGGTA    396

Qy   51 AspSerLeuProLysGluArgGlyTyrThrArgThrArgTyrLeuPheLeuGlnGlyPhe  70
           :::
Db  397 TCCAACTCTCCCGGCGCTGAGGCGGTGGTATCG-----CTCGGTGCAGTACCAGGAACTAC 450

Qy   71 TrpCysGlnAlaLysGlyIleGlnIleAlaIleMetSerPheGlnLysHisPheGlnSerLeu  90
Db  451 TGGCTGAACCCCGGCAGGCTCCAGACACATCCCGTGAAAGAGCTTCACAAGCCTCGC     510

Qy   91 GluAsnAspValValLeuAlaThrIleProLysSerGlyThrThrTrpLeuLysGluAlaLeu 110
Db  511 GCCGACGACAATCGTCTTCGGCCACCTACCTGAAAATGCGGACCACGTGGCTGGAAGGCGTG 570

Qy  111 ThrPheThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThrAsnHisPro  130
Db  571 GCCTTCGCCCATCACGACCCGGCGCCCACCCACCCAGCCCTTTCGCCCGCGGACCACCCC  630

Qy  131 LeuPheThrSerAsnProHisAspLeuValProPhePheGluTyrLysLeuTyrAlaAsn  150
Db  631 CTGCTCACCGTGCACCCGCAGGAGCCGTGCCGAGGTGCACACGCCGGGGCAG         690

Qy  151 GlyAspValProAspLeuSerGlyLeuAlaSerProArgThrPheAlaThrHisLeuPro  170
Db  691 GGC---CTGGCGGACAATCGAGAGGGCGTGCCCGTGCGCCCAGGCGTGCTCCGGACACTTGCCC 747

Qy  171 PheGlySerLeuLysGluThrHisrIleGlyLysProGlyValLysValTyrLeuCysArg  190
Db  748 CTCTCCCCGTCTCCCCGGCGCGGCCACCTCCGGCTGCCCCGTCGTTAACTTGTGCGCG     807

Qy  191 AsnProPheAspThrPheIleSerSerTrpHisTyrThrAsnAsnIleLysSerGluSer  210
Db  808 CAGCCCAAGGACGGTGTTCGTCCTCGCACTTCCTGGCACATTCGCCGCGG---GGC      864

Qy  211 ValSerProValLeuLeuAspGlnAlaPheAspLeuTyrCysArgGlyValIleGlyPhe  230
Db  865 AGGTCCCCCGTGGAGCTGGACCGCCGGCTGCCCATGTTTCTCGGAGGGGGTTCGCCTTTC  924

Qy  231 GlyProPheTrpGluHisMetLeuGlyTyrThrArgGluSerLeuLysProGluLys    250
Db  925 GGGCCCGTCTGGGAGCACTACCTCGAGTACTCGAAGGAGAGCCTGAAGCCCTGAGCGG    984

Qy  251 ValPhePheLeuArgTyrGluAspLeuLysAspAspIleGluThrAsnLeuLysArgLeu  270
Db  985 GTCCTTTTTCCTCAGGTACGAGGATGGTGGGGGCGACCCGGTCGTGCCGTGAGGACGCTC 1044

Qy  271 AlaThrPheLeuGluLeuProPheThrGluGluGluLysGlyValValLysLysAla    290
Db 1045 GCCGGCGTTCTTCGCCCGTTCACCGACGAGGAGAAGGGGGTTCCCGGCGAG         1104

Qy  291 IleAlaGluLeuCysSerPheGluAsnLeuLysGluGlyValAsnLysSerAsnLys    310
Db 1105 ATTGTGGAGGCGTGCAGCTTCGAGAATCTGAAGGAAGGCTTGGAAACGAATCCGACCGGA--- 1161

Qy  311 SerIleLysAsnPheGluAsnArgPheLeu--------------PheArgLysGlyGlu  325
Db 1161

Db 1162 GATCTTGATTGCGGAGCAACACGGTGAATCGGAAAAATCTACCTTCTTCAGGAAGGGCAAG 1221

Qy  326 ValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArgLeuSerAlaLeuVal  345
Db 1222 GTTGGGGGACTGGGAGGAACCACTTGACAAAAGAGGATGGGCAAGAAGAGGTCGACGCCGTTTC 1281

Qy  346 AspAspLysLeuGlyGlySerGlyLeuThrPhe                            356
Db 1282 GAAGACAAGCTCCAGGGGGATCGTCTCGTGTTT                           1314
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 23, 2005, 19:18:45 ; Search time 284 Seconds
656.003 Million cell updates/sec

Title: US-10-019-931-3
Perfect score: 1898
Sequence: 1 MATSSMKSIPMAIPSFSMCH...........RLSALVDDKLGGSGLTFRLS 359

Scoring table: BLOSUM62

Xgapop	10.0 ;	Xgapext	0.5
Ygapop	10.0 ;	Ygapext	0.5
Fgapop	6.0 ;	Fgapext	7.0
Delop	6.0 ;	Delext	7.0

Searched: 4168288 seqs, 25947437 residues

Total number of hits satisfying chosen parameters: 8336576

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_P/US10019931/runat_16122005_165220_18532/app_query.fasta_1.519
-DB=Published_Applications_NA_New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10019931 @CGN_1_183 @runat_16122005_165220_18532
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLX=120 -DEV TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:
Published_Applications_NA_New:*

1:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*	
2:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*	
3:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*	
4:	/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*	
5:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*	
6:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*	
7:	/cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*	
8:	/cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*	
9:	/cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*	
10:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*	

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	274	14.4	1718	7	US-11-179-843-2	Sequence 2, Appli
2	99	5.2	1027	6	US-10-750-185-42392	Sequence 42392, A
3	90	4.7	165857	7	US-11-121-086-34	Sequence 34, Appl
c 4	90	4.7	187745	7	US-11-121-086-83	Sequence 83, Appl
c 5	87.5	4.6	2412	7	US-10-821-234-791	Sequence 791, Appl
c 6	86.5	4.6	909	6	US-10-793-626-1143	Sequence 1143, Ap
c 7	86.5	4.6	3315	6	US-10-793-626-3820	Sequence 3820, Ap
c 8	86.5	4.6	3431	6	US-10-793-626-4435	Sequence 4435, Ap

c 9	85.5	4.5	1316	6	US-10-750-185-55045	Sequence 55045, A
c 10	85	4.5	171423	7	US-11-121-086-85	Sequence 85, Appli
c 11	84.5	4.5	67088	7	US-11-117-187-186	Sequence 186, App
c 12	84	4.4	420	7	US-11-121-086-42	Sequence 42, Appl
c 13	84	4.4	142303	7	US-11-821-234-822	Sequence 822, App
c 14	83.5	4.4	1281	7	US-10-750-185-31307	Sequence 31307, A
c 15	82.5	4.3	2112	7	US-11-094-586-1	Sequence 1, Appli
c 16	82.5	4.3	3312	6	US-10-793-626-3465	Sequence 3465, Ap
c 17	82	4.3	2442	6	US-10-793-626-3843	Sequence 3843, Ap
c 18	82	4.3	3084	6	US-10-793-626-3105	Sequence 3105, Ap
c 19	82	4.3	3148	6	US-10-793-626-3369	Sequence 3369, Ap
c 20	81.5	4.3	2069	7	US-11-112-908-139	Sequence 139, App
c 21	81.5	4.3	156297	7	US-11-121-086-65	Sequence 65, Appl
c 22	81	4.3	1839	6	US-10-750-185-42616	Sequence 42616, A
c 23	81	4.3	7292	7	US-11-075-646-5	Sequence 5, Appli
c 24	80.5	4.2	1169	7	US-11-112-908-414	Sequence 414, App
c 25	80.5	4.2	1549	6	US-10-995-561-18	Sequence 18, Appl
c 26	80.5	4.2	1943	7	US-11-117-187-204	Sequence 204, App
c 27	80.5	4.2	109974	7	US-11-117-187-57	Sequence 57, Appl
c 28	80.5	4.2	161994	7	US-11-112-908-57	Sequence 59, Appl
c 29	80.5	4.2	168656	7	US-11-112-908-59	Sequence 58, Appl
c 30	80.5	4.2	170285	7	US-11-112-908-58	Sequence 58, Appl
c 31	80	4.2	2250	7	US-11-112-908-137	Sequence 137, App
c 32	80	4.2	2705	6	US-10-793-626-50240	Sequence 50240, A
c 33	80	4.2	2950	6	US-10-793-626-3349	Sequence 3349, Ap
c 34	80	4.2	4198	6	US-10-793-626-3604	Sequence 3604, Ap
c 35	80	4.2	1080000	6	US-10-928-446A-1	Sequence 1, Appli
c 36	80	4.2	1080000	6	US-10-928-446A-181	Sequence 181, App
c 37	80	4.2	1080000	6	US-10-928-446A-183	Sequence 183, App
c 38	80	4.2	1080000	6	US-10-928-446A-185	Sequence 185, App
c 39	80	4.2	1080000	6	US-10-928-446A-187	Sequence 187, App
c 40	80	4.2	1080000	6	US-10-928-446A-189	Sequence 189, App
c 41	80	4.2	1080000	6	US-10-928-446A-191	Sequence 191, App
c 42	80	4.2	1080000	6	US-10-928-446A-193	Sequence 193, App
c 43	80	4.2	1080000	6	US-10-928-446A-195	Sequence 195, App
c 44	80	4.2	1080000	6	US-10-928-446A-197	Sequence 197, App
c 45	80	4.2	1080000	6	US-10-928-446A-199	Sequence 199, App

ALIGNMENTS

RESULT 1
US-11-179-843-2
Sequence 2, Application US/11179843
Publication No. US20050277153A1
GENERAL INFORMATION:

; APPLICANT: Thomae, Bianca A.
; APPLICANT: Wieben, Eric D.
; APPLICANT: Weinshilboum, Richard M.
; TITLE OF INVENTION: Sulfotransferase SULT2A1 Sequence
; TITLE OF INVENTION: Variants
; FILE REFERENCE: 07039-276001
; CURRENT APPLICATION NUMBER: US/11/179,843
; CURRENT FILING DATE: 2005-07-12
; PRIOR APPLICATION NUMBER: US/10/177,570
; PRIOR APPLICATION NUMBER: US 60/300,165
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/300,165
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/300,165
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 2
; LENGTH: 1718
; TYPE: DNA
; ORGANISM: Homo Sapiens

US-11-179-843-2
; Query Match:

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:

Pred. No.:	1.12e-21
Score:	274.00
Percent Similarity:	46.24%
Best Local Similarity:	27.24%
Query Match:	14.44%
	7

Length:	1718
Matches:	76
Conservative:	53
Mismatches:	118
Indels:	32
Gaps:	9

US-10-019-931-3 (1-359) x US-11-179-843-2 (1-1718)

```
Qy   78  GlnAlaIleMetSerPheGlnLysHisPheGlnSerLeuGluAsnAspValLeuAla   97
             ::|          |||||||||::::
Db   81  GaaAccTTaAgaAaagTacgTgaTgagTTcgTgaTaAggaTgaagaTgTaaTaaTTg  140

Qy   98  ThrIleProLysSerGlyThrThrThrLeuLysAlaLeuThrPheThrIleLeuAsnArg  117
             :::       |||:||:::
Db  141  ACTTACCCaAaaTcAggaAcaAaaACTggTTggcTgagATTcTgcTgacTcCaAg    200

Qy  118  HisArgPheAspProValAlaSerSerThrArgAsnHisProLeuPheThrSerAsnProHis  137
             :::
Db  201  GGGGaTgccaAgTggaTccaaTcT                                    224

Qy  138  AspLeuValProPheGluThrTyrLysLeuThrTyrAlaAsnGlyAspVal-----ProAsp  155
             |||:||                          :::       :::
Db  225  ------gTgccCaTcTgggAgccaTcaccccTgggTagggagagTgTaTacAgca      278

Qy  156  LeuSerGlyLeuValSerProAsnArgThrPheAlaThrHisLeuProPheGlySerLeuLys  175
             :::|||            ::::::|||
Db  279  cTcaGTgaAaAcggaGagTccacGTTTaTTccTcccAcccccaTcccAgTTaTTccc      338
```

US-10-019-931-3 (1-359) x US-10-750-185-42392 (1-1027)

```
Qy  176  GluThrIleGlyLysProGlyValAlaTyrValValTyrLeuCysArgAsnProPheAspThr  195
             |||            |||:::|||
Db  339  AAGTCTTTcTTcagTTccAAggCCaAgGTgAaTTaTcTcaTgaGaaAaTCCCaGAgAaTgTT   398

Qy  196  PheIleSerSerTrpHisTyrThrArgAsnAsnIleLeuLysSerGluSerValSerProValLeu  215
             :::|||                  ::::::|||
Db  399  TTGGGTGTCTCGgTTgTTaTTTTTcTTcggAAaaaAcaTgAAGTTTaTTaAG---AAaCAAaGTca  455

Qy  216  LeuAspGlnAlaPheAspLeuThrTyrCysArgGlyValIleGlyPheGlyValValIleGlyTyr  235
             :::::::::|||                        :::::::::
Db  456  TGGGaGGAaAAaTaTTTTgAAaTggTTTTTgTcAAGgaaAcgTgTcGTaTaTgggTgTTTgAc     515

Qy  236  HisMetLeuGlyTyrTyrTrpArgGluSerLeuLysLysValPhe---PheLeu              254
             |||            :::
Db  516  CACaTTCaTggC----TGGaTcCCC---------AtgAGagAGgAgaAaAaACTTCCTgTTACTg   566

Qy  255  ArgThrTyrGluAspLeuLysAspIleGluThrArgAsnLeuThrArgPheLeu             274
             ::::::|||||||||                 :::
Db  567  AGTTaTcGaGGacTcagAAacACAggAgaAGaAGAgAaTcCTgcAacTTcaaTTCCTg        626

Qy  275  GlyLeuProPheThrGluLeuGlyLeuGluArgLysGlyValValLysValAlaIleAlaGluLeu  294
             :::|||
Db  627  GGAAAgAcgTTaGAaccCGAGaA------cTgAAacTTaATTcTcaaGaAaC              671

Qy  295  CysSerPheGluAsnLeuLysGlyLeuValAlaAsnLysSerAsnLysSerIleLeuLysAsn    314
             |||
Db  672  AGcTcCTTTcaGAgAcAaTgAAa---GAAacaAgaTgTCcAaTTaTTccTcCTgTacagT      725

Qy  315  ---------PheGluAsnArgPheLeuGluArgLysSerGlyLeuValSerAspTrpVal       330
             :::
Db  726  GTTGaTTaGTTaGTggGACaAAaGCACaAcaCTTcTcGagAaAagGGTaTcTGGGaGCTgGGAAa  785

Qy  331  AsnTyrLeuSerProSerGlnValGlnArgLeuSerAlaLeuValAspAspLysLysLeu      349
             :::
Db  786  AATcACTTcAcAgTGCCCAagCTGAAgACTTTgATaAaTTGTTcAAgAGAagAGaTG         842
```

Alignment Scores:
```
Pred. No.:
Score:                    0.294
Percent Similarity:       99.00
Best Local Similarity:    35.98%
Query Match:              21.96%
DB:                       5.22%
```

```
Length:        1027
Matches:         47
Conservative:    30
Mismatches:      76
Indels:          62
Gaps:             7
```

US-10-019-931-3 (1-359) x US-10-750-185-42392 (1-1027)

```
Qy  123  ValAlaSerSerThrArgAsnHisProLeuPheThrSerAsnProHisAspLeuValPro  142
             :::|||                        :::
Db  274  CTgGCaCaGaTTcTgGgGTTcAaaAcAaTTTcTTcaAAgTCTTCaCAgTgcTgGCTgT      333

Qy  143  PheGluThrTyrLysLeuThrTyrAlaAsnGlyAspValProAspLeuSerGlyLeuValAlaSerPro  162
             |||:::
Db  334  TTTgCTTcTTcCCTTTcCTgAAC-----------------------------------------------   357

Qy  163  ArgThrPheAlaThrHisLeuProPheGlySerLeuLysGluThrIleGlyLysProGly  182
             :::|||
Db  358  --ACTTTACTTcTTcACcTCCCACTc---------------------------------------      381

Qy  183  ValLysValValTyrLeuCysArgAsnProPheAspThrPheIleSerSerTrpHisTyr  202
             :::|||
Db  382  ----ACCCcTTTTaAaTAtgcTTTTaTcTTgCTggGCAACAT                         417

Qy  203  ThrArgAsnAsnIleLeuLysSerGluSerValSerProValLeuLeuLeuAspGln-----  218
             |||:::
Db  418  -----TcTTaTCaAaAGTcAAcACaGTaAAaCTcCTaCTAACCaAgAaAaAATgGGCCTTCAGGa   471

Qy  219  AlaPheAspLeuThrTyr--------------CysArgGlyValIleGlyPheGlyGlyPhe  233
             |||:::
Db  472  CTGTTTTTaTcTaTTTgACTTcCCcTATaAaTGTTGGaGGcTTcCgTTgTTCCTTTC         531

Qy  234  TrpGluHisMetLeuGlyTyrTyrTrpArgGluSerLeuLysSerArgProGlyLysValPhePhe  253
             |||:::
Db  532  TGGAaACACTcT----------------------------------------------------------TTT  546

Qy  254  LeuArgThrTyrGluAspLeuLysAspIleGluThrArgAsnLeuThrArgLeuAlaThrPhe  273
             :::
Db  547  CTTTTcTTcTGGGATTcTcGcTCCcTcCcTTTTcTGaTcATTTCTTcCATcTcTACcTTc      606

Qy  274  LeuGluLeuProPheThrGluLeuGlyLeuGluArgLysGlyValValLysValAlaIleAlaGlu  293
             |||:::
Db  607  CTTT---------TTcTgTcCTGaAAAaTgTTGACAAGTTCCCAGTTC--TcTgTcTTAGGc      655

Qy  294  LeuCysSerPheGluAsnLeuLysGlyLeuValAlaAsnLysSerAsnLysSerIleLeuLys  313
             :::|||
Db  656  CTTTTgTCTAAaTCTTcTGgGCTcCCcTcCAGTGAGTGGCAGaATACTgCTgGTcGTCaTGCGAAG  715
```

APPLICANT: MMI GENOMICS, INC.
APPLICANT: DeNISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis

TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MMI1100-2

APPLICANT: POULSEN, TIM S.
APPLICANT: NIELSEN, KIRSTEN V.

```
Qy  314  AsnPheGluAsnArgPheLeuGluArgLysSerGlyLeuGluValSer  327
             :::
Db  656  CTTTTgTCTAAaTCTTcTGgGCTcCCcTcCAGTGAGTGGCAGaATACTgCTgGTcGTCaTGCGAAG  715

Qy  294  LeuCysSerPheGluAsnLeuLysGlyLeuValAlaAsnLysSerAsnLysSerIleLeuLys  313
             |||
Db  607  CTT---------TTcTgTcCTGaAAAaTgTTGACAAGTTCCCAGTTC--TcTgTcTTAGGc      655

Db  716  AATTATTcTTTTTcCTTcTTGTTcTgGTTcAcGGAAAAgGGTGAcCTCTCC  757
```

```
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138.6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentin version 3.3
SEQ ID NO 34
TYPE: DNA
ORGANISM: Homo sapiens
LENGTH: 165857
US-11-121-086-34

Alignment Scores:
Pred. No.:
Score:                    9.17e+03
Percent Similarity:       90.00
Best Local Similarity:    30.82%
Query Match:              22.66%
DB:                       4.74%

                          Length:          165857
                          Matches:         75
                          Conservative:    27
                          Mismatches:      95
                          Indels:          134
                          Gaps:            18

US-10-019-931-3 (1-359) x US-11-121-086-34 (1-165857)

Qy   26  LysGluGlyLysThrArgAspValProLysAlaGluValGlyLeuSerCysGlu   45
         |||||||
Db 8825  AAGACAGGGTTGTGTCATTTTCTACCAAATTCACAGCAACATGAATGGCTTTTTCAAA 8884

Qy   46  PheGlnGluMetLeuAspSerLeuProLysGluArgGlyThrArgTyrLeuTyr   65
         |||||||   ::::
Db 8885  TTTCAG------TTATTTGAGCTTCCTCAACAATTAGGGTGGGAGA 8923

Qy   66  LeuPheGlyGlyPheTrpCysGlnAlaLysGluGlnIleAlaIleMetSerPheGlnLys   85
         |||||||   :::
Db 8924  ---TACACTAACATTTGGTTAACAACAAAACAACCAATCCCCTCCCCCCAAAAA 8980

Qy   86  HisPheGlnSerLeuGluAsnAspValValLeuAlaThrIlePro------LysSerGly 103
         |||||||
Db 8981  TAATTT---------TATTTAGAAAATTAACCTCCCTAAAAATCTGGC 9019

Qy  104  ThrThrTrp-----   106
         ||||
Db 9020  ATGTTTTGAGGCAAAGAGAGTAATGAATAAAAACACAAAACAAACAAGAAACT 9079

Qy  121  ---------------AspProValAlaSer---ThrAsnHisProLeu  131
                        :::
Db 9080  TTCCATGTAGCCAAAGTTCTCACAACACCCATCCTGAAATGCACTACTTTATCTCAGG 9139

Qy  107  -------LeuLysAlaLeuThrPheThrIleLeuAsnArgHisArgPhe------- 120
               |||||||
Db 9140  AGACAAGTAAGGTAAACTATTATCAAGGCCATAGGGAAGCTCCAAAAAAAATCATGAAGTA 9199

Qy  132  PheThrSerAsnProHis-------------------AspLeuValProPhePheGlu  144
         |||||||                              ||:::
Db 9200  CAAAGTTCAAACAACAATACAGGTTTTGAAGGGCTTTGACAAGCTTCCACCTTCTTACCTA 9259

Qy  145  TyrLysLeuTyrArgLeuAsnGlyAspValPro------   154
         :::||||||
Db 9260  TGGGAGACTTGTTAGCTTAACACAAATGTCCCATTATACTGCGTTTTTCCAGCAGCTAA 9319

Qy  155  ---------------AspLeuSerGlyLeuAlaSerProArgThrPhe 165
                        ||||||
Db 9320  AACCTGTGAAACATGTTAAAAAAGAAAAAAATTAAGTGGGGTGGGG--- 9367

Qy  166  AlaThrHisLeuProPheGlySerLeuLysGlyThr------ 177
         |||
Db 9368  TTGAGGTAGATCCAATTCAGCAGTACCACAGAAACTAAAGTTTAAATAGGAGGACTCTCA 9427

Qy  178  ---IleGluLysProGlyValLysValValTyrLeuCysArgAsnProPheAsp----- 194
            :::::::|||
Db 9428  AAACTCCAGCAGCCTAGGGTTCCCAGCAGCAATGACAAAATCCAAG---CCTTTTTCATTCCTG 9484

Qy  195  -------------ThrPheIleSerSer---TrpHisTyrThrAsnAsnIleLysSerGluSer 210
                                 |||
Db   9692  CCAAAGAACAAAAAGAAAGAGAAAAAATCTTTTTT 9724
```

```
Db 9485  TTGCTTCCCACCATTCAATCTCTTTCCCTTTGG------------ 9514
         ||||||||||||                              |||
Qy  211  ValSerProValLeuLeuValAspGlnAlaPheAspLeuTyrCysArgGlyValValIleGlyPhe 230
         |||:::||||||
Db 9515  ---CTTCCAAGAACATCACATACTGCTTACAACCTTGACTGGGCCTCAGCATACCATAAA 9571

Qy  231  GlyProPheTrp------------GluHisMetLeuGlyTyrTrp---ArgGluSerLeu--- 245
         |||
Db 9572  GGTCCATTGTACTCAATTTTAGAGTTTTTTTTGGGGGTGGGGAGGAATTCTTTTTTG 9631

Qy  245  ---------------------------------LysArgProGluLysValPhePhe 253
         :::
Db 9632  AGCAGTTCTAACTATATATATATAAGGATTGTATTTTGTTGTTTTTTCTTGGACACCTT 9691
```

```
RESULT 4
US-11-121-086-83/c
Sequence 83, Application US/11121086
Publication No. US20050266459A1
GENERAL INFORMATION:
APPLICANT: NIELSEN, TIM S.
APPLICANT: POULSEN, KIRSTEN V.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138.6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patentin version 3.3
SEQ ID NO 83
TYPE: DNA
LENGTH: 187745
ORGANISM: Homo sapiens
US-11-121-086-83

Alignment Scores:
Pred. No.:
Score:                    1.11e+04
Percent Similarity:       90.00
Best Local Similarity:    34.45%
Query Match:              21.07%
DB:                       4.74%

                          Length:          187745
                          Matches:         63
                          Conservative:    40
                          Mismatches:      113
                          Indels:          83
                          Gaps:            11

US-10-019-931-3 (1-359) x US-11-121-086-83 (1-187745)

Qy   64  LeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGluIleGlnAlaIleMetSerPhe  83
         |||:::||||||:::||||||
Db 123825 CTCTTCCTTCCTTCAGGGGCAGAAAACGCCCAAGCAGGC------------ 123790

Qy   84  GlnLysHisPheGlnSerLeuGluAsnAspValValLeuAlaThrIleProLysSerGly 103
         :::|||||::::||
Db 123789 ----------------CTCTTCCACCTTTACCACCCACTGGG 123763

Qy  104  ThrThr------------TrpLeuLysAlaLeuThrPhe------ 112
         |||
Db 123762 GTCACAGCATCCTTTCTCTGCTGTGGAGGGGTTGGGCGGTTGGGGCCACAAAACCCACCCCTGCCC 123703

Qy  113  ------ThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThrAsnHisPro 130
               :::|||||||
Db 123702 TTCCCCAGCCTTGCCAGGCAGCACAGATCAGCTGATTGCCTCGCCCTCATCTGGCATTCC 123643

Qy  131  LeuPheThrSerAsnProHisAspLeuValProPhePheGluTyrLys-LeuTyrArgLeuAs 150
         ||||||:::|||
Db 123642 AAGCACCACGGGACTACTGGCAGCACAGTGTCTTGGGCCGGAGGCCTTGAGGCCAC 123583

Qy  150  nGlyAspValProAspLeuSerGlyLeuAlaSerProArgThrPheAlaThrHisLeuPr 170
         |||:::
Db 123582 TGCCTCC---------CCAGGGCTCCTCCTGAAGCCAAAGAATGAAGTGGGTGAGCTGAG 123535
```

```
Qy    170 oPheGlySerLeuLysGluThrIleGlyValLysProGlyValValValTyrLeuCysAr 190
          |||        :::                    |||||||||||||||  :::
Db 123534 CTACGGAAGAGTGATGGGTCCCTTGGAAGACCCTGGTGGAGG------CACAAATGCCG 123481

Qy    190 gAsnProPheAspThrPheIleSerSerTrpHisTyrThrAsnIleLysSerGlue 210
          |||||
Db 123480 CTTCTGTGCCAAAGTATTTGGCAGT------------------------------ 123456

Qy    210 rValSerProValLeuLeuAspGlnAlaPheAspLeuTyrCysArgGlyValIleGlyPh 230
          |||
Db 123455 ------GACAGTGCCCTGCAGATCCAACCTTCGTTCCACAGGGGTGA--- 123415

Qy    230 eGlyProThrTrpGluHisMetLeuGlyTyrTrpArgGluSerLeuLysArgProGluLy 250
          |||||                                         |||
Db 123414 GAGGCCCTATAAGTGCAAATGTCTGTGGAAAC---CGTTTTACCACCGTTGGCAACTTCAA 123358

Qy    250 sValPhePheLeuArgThrGluAlaTyrGluAspGlnSerAsnIleGluThrAsn----- 266
          |||                                         |||
Db 123357 AGTGCANTTCCACCCGCATCGTGAAGAAGTACCACATGTGCAGAGTCANGAACCCACACCCAGT 123298

Qy    267 ------LeuLysArgLeuAlaThrPheLeuGluLeuProPhe--------------- 278
                |||                                     :::
Db 123297 ACCAGACCACTTAGACTATGTCAATTACCAGCAGTGGCTTGGCCTTATGGTATGTCCGTGCC 123238

Qy    279 --------ThrGluGluGlu--------------GluArgLysGlyVa 287
                  |||||||                    :::
Db 123237 ACCAGAGAAGCCGAGGAGGAGGAGGAGGGGGTTGAGCGCAAGCCTCT 123178

Qy    287                                              123118
          |||
Db 123177 GGTGGCGCTCCACAACAGCACTCAGTGCCACAGAGGACGCCTGACTCGCTCTCCACCAGTGC 123118

Qy    307 sSerAsnLysSerIleLysAsnPheGlu------AsnArgPheLeuPheArgLys 323
          :::                       |||||||||
Db 123117 AGGCACAGCCAGGACTCCCAGGACTCCCTGCTTTCAATAAGTTTGTGCTCATGAAA 123063
```

RESULT 5

US-10-821-234-791

; Sequence 791, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 791
; LENGTH: 2412
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-821-234-791

Alignment Scores:

Pred. No.:	
Score:	25.6
Percent Similarity:	87.50
Best Local Similarity:	34.70%
Query Match:	20.96%
	4.61%
	6

Length:	2412
Matches:	87
Conservative:	57
Mismatches:	138
Indels:	133
Gaps:	24

```
Qy     43 SerCysGluPheGlnLeuGlnMetLeuTrpAspSerLeuProLysGluArgGlyGly------ 58
          :::                                       ||||||
Db    931 GCAGTAGAG---GAAGAAGAAGAAAGAAAAGAAAGACTAAAAAAAGTTGAAAAAACT 987

Qy     59 ---TrpArgThrArgTyrLeuTyrLeuPheGlnIn GlyPheTrpCysGln--AlaLysGlu 76
             :::                           |||
Db    988 GTCTGGGACTGGGAACTTATGAATGATAATCAAAACCAATATGGCAGAGACCATCAAAAAGAA 1047

Qy     77 IleGln----AlaIleMetSerPheGlnInLysHisArgPheGlnSerLeuGluAspVal 94
          |||                           |||
Db   1048 GTAGAAGAAGGATGAAATACAAAAGCTTTCTTACAAAATCATTT---TCAAAGGAAAGATGAC 1104

Qy     95 ValLeuAlaThrIle------ProLysSerGlyThrThrThrGlnLeuLeuAlaLeuThrPhe 112
          |||                |||:::                                  |||
Db   1105 CCCATAGCTTATATTCACTTTACTGCTGAAGGGGAAGTTACCTTCAAATCAATTTATT 1164

Qy    113 -----ThrIleLeuAsnArgHisArgPheAspProValAlaSerSer----- 126
               |||                                       |||
Db   1165 GTACCCACATCTCAGTCCAGTGGTCGTGTTGACGGAATTATGGATCTAAAAGAGACGGATTAC 1224

Qy    127 ------ThrAsnHisProLeuPheThrSerAsnProHisAspLeuValProPhe 142
                :::                                         ||| 
Db   1225 ATTAAGCTCTATGTCGCGCGTGTTATCATCACAGAGACGACTTCCATGAATATGAGNGCCT--- 1281

Qy    143 PheGluTyrLysLeuTyrArgAlaAsnGly------AspValProAspLeuSerGlyLeuAla 160
          |||                           :::
Db   1282 ---AAAATCCCTAAATTTGTCAAGGGTGTGGGTGGAGTCTCCCCTTGAAATGTT 1338

Qy    161 SerProArgThrPheAlaThrHisLeuProPheGlySerLeuValGluThrIleGluLys 180
          |||                                                      |||
Db   1339 TCCCGGCGAGACTTCAGCAACAT---AAACTGCTTAAGGTGTTAGAGAAG 1386

Qy    181 ProGlyValLys------ValValTyrLeuCysArgArgAsnProPhe--Asp 194
          :::                                       |||
Db   1387 AAGCTTGTTCGTAAAAGCTGGACANTCAAGAGAAGATTGCTGATGATAATACAATGAT 1446

Qy    195 ThrPhe 196
          |||
Db   1447 ACTTTTTGGAAAGAATTTGGTACCAACATCAAGCTTCGGTTGATTGAAGAGACCACTCGAAT 1506

Qy    197 ------IleSerSerTrpHisTyrThrAsnIleLys 207
                |||
Db   1507 CGAACACGTCTTGCTAAACTTCTTAGGTTCCAGTCTTCTTCATCACTGACACATTACT 1566

Qy    208 SerGluSerValSerProValLeuLeuAspGlnAlaPheAspLeuTyrCysArgGlyVal 227
          |||
Db   1567 AGC------CTAGACCAG--- 1578

Qy    228 IleGlyPheGlyProPheTrpGluHisMetLeuGlyTyrTrpArgGluSerLeuLysArg 247
          :::
Db   1579 ---TATGTGGAAAGAAATGAAGGAAAA 1602

Qy    248 ProGluLysValPhePheLeuArgThrGluAlaTyrGluAspGlnSerAsnIleGluThrAsnLeu 267
          |||                                            :::
Db   1603 CAAGACAAAATCTACTTCATGGCTGGGGTCCAGCAGAAA---GAGGCTGAACTTCT--- 1656

Qy    268 LysArgLeuAlaThrPheLeuGluLeuProPheThrGluGluGluGluGluArgGlyGly 286
          :::                                       ||||||||
Db   1657 ---CCATTTGTTGAGGCGACTTCTGAAAAAGGGCTAT 1689

Qy    287 ------ValValLysAlaIleAlaGluLeuCys--- 295
                :::                     |||
Db   1690 GAAGTTATTTACCTCACAGAACCTGTGCGATGAATACTGTATTCAGGCCGCCCTTCCCGAATT 1749

Qy    296 -----SerPheGluAsnLeuLysGlyLeuValAsnLysSerAsnLysSerIle 312
               |||:::                                       |||
Db   1750 GATGGGAAGAGAGGTTCCAGAATGTTGCCAAG---GAAGGAGTGAAGTTCGATAAAAGTCAG 1806

Qy    313 LysAsnPheGluAsnArgPheLeuPheArgLysGlyGluValValSerAspTrpValAsnTyr 332
          |||                                                      :::
Db   1807 AAAACTAAGGAGGAGTGGTTGATTAGAAAGGCAGTTGAAGCCTCGCTGAATTGG 1863
```

US-10-019-931-3 (1-359) x US-10-821-234-791 (1-2412)

```
Qy     23 GluLeuLeuLysGlyLysThrArgAspAspValProLysAlaGlyGluGluGluGlyTyrLeu 42
          :::                       ||||||                  :::
Db    871 GAGCCCATGGAGGAAGAGGAAGAAAGAGAAGAGAAACTTGATGATGAAGCT 930
```

RESULT 6 (left column)

```
Qy  333 LeuSerProSerGlnVal------------------------------GluArgLeuSer 342
Db 1864 ATGAAAGATAAAGCCCTTAAGGACAAGATGATGAAAAGGCTGTGGTTGTCTCAGCCGCCTGACA 1923

Qy  343 ------AlaLeuValAspAlaGlyLeuGlyGlySerGly 353
Db 1924 GAATCTCCGTGTGTGCTTGGTCGTGGCCAGCCAGTACGGATGAGGTCTCGGC 1968
```

RESULT 6
US-10-793-626-1143
; Sequence 1143, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1143
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-1143

Alignment Scores:
Pred. No.:
Score: 7.39
Percent Similarity: 86.50
Best Local Similarity: 37.79%
Query Match: 22.41%
                        4.56%
                        6
Length: 909
Matches: 67
Conservative: 46
Mismatches: 87
Indels: 99
Gaps: 18

US-10-019-931-3 (1-359) x US-10-793-626-1143 (1-909)

Qy   10 ProMetAlaIleProSerPheSerMetCysHisLysLeuGluLeuLeuLysGluGlyLys 29
Db 130 CCTGTACTATTACCTAAATTT-----------GTAGAATTGATGGCTCAAGGAAAT 174

Qy   30 ThrArgAspValProLysAlaGluGluAspGlu---GlyLeuSerCysGluPheGlnGlu 48
Db 175 ---------CCTTCATTGCAAAAAGTAGCAATAGAATTAAGAGGTAAGGAA 222

Qy   49 MetLeuAspSerLeuProLysGluArgGlyTrpArgThrArgTyrLeuTyrLeuPheGln 68
Db 223 TTATTTAAAATCTCTA---------- 237

Qy   69 GlyPheTrpCysGlnAlaLysGluIleGlnAlaIleMetSerPheGlnLysHisPheGln 88
Db 238 --------GTTAAAGAAATTGATATTATTAAGT 264

Qy   89 SerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThrTrpLeuLys 108
Db 265 ----------CGTAATGACGTGATTTTAGATGTTATG 300

Qy 109 AlaLeuThrPheThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThrAsn 128
Db 301 GCTAAAACTGAAAAAAATTATTTCACGTTATAAA 348

Qy 129 HisProLeuPheThrSerAsnProHisAspLeuValProPheGluTyrTyrLysLeuTyr 148
Db 349 CACCCACTAAATGAACAAGGACAAAGATCATTAATAAGATGTCCAATA 393

Qy 149 AlaAsnGlyAspValProAspLeuSerGlyLeuAlaSerProArgThrPhe------ 165
Db 394 ---AATGAGGAAGCTGAACATCTACCAGAAATCCCGAAGTACATCGGCTAATTACGTTGATTGAT 450
```

## RESULT 7 (right column)

```
Qy  166 AlaThrHisLeuProPheGlySerLeuLysLysGluThrIleGluGluLysProGlyValLysLysVal 185
Db  451 ACAACAAAAATTAAAACCCTAAAGAATTAAAAGCAACCGCATTTCAAAG-------- 495

Qy  186 ValTyrLeuCysArgGlyAsnProPheAspThrPhe------IleSerSerPheHisTyr 202
Db  496 TTTTATTTAGATGAAAAC---TTTGAAACATTTACAATACAAATCAACGTGACAAGTTCGGTTTC 552

Qy  203 ThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnAlaPheAspLeu 222
Db  553 AAGCAATGGTATACAAATGGATTGCT-----------GATTTAGTTGTTTTTGATGTC 594

Qy  223 ----------TyrCysArgGlyValIleGlyLysGlyProGlyTrpGlu 235
Db  595 AGATTTCTACCTAATCCCTACTATGTAGAGGGAATTGCGTCCATTTACGTTAGATGAG 654

Qy  236 HisMetLeuGlyTyr------------TrpArgGluSerLeuLysArgProGlyLysLys 252
Db  655 CCAGTGTACAATTACGTTATGAAGTGGGAAAGAAACC-------- 699

Qy  253 PheLeuArgTyrGluAspPheLeuLysAspIleGluThrAsnLeuLysArgGlnLeuAlaThr 272
Db  700 TTTGATGTAAATTAAACAGGATTTA---------- 732

Qy  273 PheLeuGluLeuProPheThrGluGluGluGluArgLysGlyValValLysAlaIle 291
Db  733 ----------AATTCCTGGCTACAAAAAAGAGAGGTAAATCGCAAATTGGTTAATTGCTAATA 780
```

RESULT 7
US-10-793-626-3820
; Sequence 3820, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3820
; LENGTH: 3315
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-3820

Alignment Scores:
Pred. No.:
Score: 55
Percent Similarity: 86.50
Best Local Similarity: 37.79%
Query Match: 22.41%
                        4.56%
                        6
Length: 3315
Matches: 67
Conservative: 46
Mismatches: 87
Indels: 99
Gaps: 18

US-10-019-931-3 (1-359) x US-10-793-626-3820 (1-3315)

Qy   10 ProMetAlaIleProSerPheSerMetCysHisLysLeuGluLeuLeuLysGluGlyLys 29
Db 540 CCTGTACTATTACCTAAATTT-----------GTAGAATTGATGGCTCAAGGAAAT 584

Qy   30 ThrArgAspValProLysAlaGluGluAspGlu---GlyLeuSerCysGluPheGlnGlu 48
Db 585 ---------CCTTCATTGCAAAAAGTAGCAATAGAATTAAGAGGTAAGGAA 632

Qy   49 MetLeuAspSerLeuProLysGluArgGlyTrpArgThrArgTyrLeuTyrLeuPheGln 68
Db 633 TTATTTAAAATCTCTA---------- 647
```

```
Qy    69 GlyPheTrpCysGlnAlaLysGluIleGlnAlaIleMetSerPheGlnLysHisPheGln 88
             |||  ||||                                ||||
Db   648 -------------GTTAAGAAATTGATATTATTAAAAGT------------------- 674

Qy    89 SerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThrTrpLeuLys 108
             :::||          |||||||||||::::
Db   675 -CGTAATGACGTGATTTTAGANTGTATG--                         ::: TTTTTAGA 710

Qy   109 AlaLeuThrPheThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThrAsn 128
          |||  ||                          ||||
Db   711 GCTAAAAATCGAAAAAATTATTTCACGTTATAA--      GAATCAAGAAGAGCG 758

Qy   129 HisProLeuPheThrSerAsnProHisAspPhePheGluTyrTyrLysLeuTyr 148
          ||:|||                              |||||||
Db   759 CACCCACTAAATGACAACAAGGACAAAGATCATTAATAGATGCAATA---     803

Qy   149 AlaAsnGlyAspValProAspLeuSerGlyLeuAlaSerProArgThrPhe--- 165
          |||  :::                               :::
Db   804 --AATGAGGAAGCTGAACAACTATCAGAAAATCCGAGTATCGCTAATTACGTGATTGAT 860

Qy   166 AlaThrHisLeuProPheGlySerLeuLysGluThrIleGluGlyValValLysVal 185
          |||  |||                                 ||||
Db   861 ACAACAAAAATTAAAACCTAAAGAATTAAAGCAACGCATTTCAAAG           905

Qy   186 ValTyrLeuCysArgAsnProPheAspThrPhe-------IleSerSerTrpHisTyr 202
          ||||||                                  :::||||
Db   906 TTTTTATTTAGANTGAAAAC---TTTGAAACATTTACAATCAACGTGACAAGTTTCGGTTTC 962

Qy   203 ThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnIleAlaPheAspLeu 222
          |||  |||                                  |||
Db   963 AAGCATAGGTATACAAATGGATGCT-                 GATTTAGTTTTTGANGTC 1004

Qy   223 -------------------TyrCysArgArgGlyValIleGlyIlePheGlyProPheTrpGlu 235
                             :::|||
Db  1005 AGATTTCTTACCTAATCCCTACTATGTRAGAGGAATTCCGTCCCATTTACTGGTTTAGNATGAG 1064

Qy   236 HisMetLeuGlyTyr---------TrpArgGluSerLeuLysArgProGlyLysValPhe 252
          |||                      :::||
Db  1065 CCAGTGTACAARTTACGTTATGAAGTGGAAAGAAACC-         CAAATATTT 1109

Qy   253 PheLeuArgGlyTyrGluAspProLeuLysAspIleGluThrArgAsnLeuLysArgLeuAlaThr 272
          |||  ||||                            |||
Db  1110 TTTGATAAATTAACAGATTTA-                GATTTAGTTTTTGANGTC 1142

Qy   273 PheLeuGluLeuProPheThrGluGlyGluGlyArgLysGlyValValLysAlaIle 291
          |||  :::                              :::||
Db  1143 -ATTCCTGGCTACAAAAAAGAAGGTAAATCGCAATTGGTTATTGCTATA 1190
```

```
Score:                     86.50      Matches:          67
Percent Similarity:        37.79%     Conservative:     46
Best Local Similarity:     22.41%     Mismatches:       87
Query Match:               4.56%      Indels:           99
DB:                        6          Gaps:             18

US-10-019-931-3 (1-359) x US-10-793-626-4435 (1-3431)

Qy    10 ProMetAlaIleProSerPheSerMetCysHisHisLysLeuGluLeuLeuGluGlyLys 29
             |||:||                                     |||
Db  1598 CCTGTACTATTACCTAAAATT--          GTAGAATTTGATGAGTCCAAGGAAAT 1554

Qy    30 ThrArgAspValProLysGluGluAspGlu---GlyLeuSerCysGluPheGlnGlu 48
          |||                               :::   |||
Db  1553 -------------CCTTCATTGCAAAAAGTAGCAATAGATTTAAGAGGTTAAGGAA 1506

Qy    49 MetLeuAspSerLeuProLysGluArgGlyTrpArgThrArgTyrLeuTyrLeuPheGln 68
          |||                                |||
Db  1505 TTATTTAAATCTCTA-            GAATCAAGAAGAGCG 1491

Qy    69 GlyPheTrpCysGlnAlaLysGluIleGlnAlaIleMetSerPheGlnLysHisPheGln 88
          |||                                   |||
Db  1490 -GTTAAGAAATTGATATTATTAAAAGT-           1464

Qy    89 SerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThrTrpLeuLys 108
          :::|||                                 |||
Db  1463 --CGTAATGACGTGATTTTAGANTGTATG-           TTTTTAGAA 1428

Qy   109 AlaLeuThrPheThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThrAsn 128
          |||:::                                 |||
Db  1427 GCTAAAAATCGAAAAAATTATTTCACGTTATAAA-     GAATCAAGAAGAGCG 1380

Qy   129 HisProLeuPheThrSerAsnProHisAspPhePheGluTyrTyrLysLeuTyr 148
          |||                                   |||
Db  1379 CACCCACTAAATGACAACAAGGACAAAGATCATTAATAGATGCAATA--     1335

Qy   149 AlaAsnGlyAspValProAspLeuSerGlyLeuAlaSerProArgThrPhe--- 165
          |||:::                                 :::
Db  1334 --AATGAGGAAGCTGAACAACATCTATCAGAAAATCCGAAGTATCGCTAATTACGTGATTGAT 1278

Qy   166 AlaThrHisLeuProPheGlySerLeuLysGluThrIleGluGlyValValLysVal 185
          |||                                   :::||
Db  1277 ACAACAAAAATTAAAACCTAAAGAATTAAAGCAACGCATTTCAAAG 1233

Qy   186 ValTyrLeuCysArgAsnProPheAspThrPhe-------IleSerSerTrpHisTyr 202
          |||:::                                 :::||
Db  1232 TTTTTATTTAGANTGAAAAC---TTTGAAACATTTACAATCAACGTGACAAGTTTCGGTTTC 1176

Qy   203 ThrAsnAsnIleLysSerGluSerValSerProValLeuLeuAspGlnIleAlaPheAspLeu 222
          |||:::                                 |||
Db  1175 AAGCATAGGTATACAAATGGATGCT-             GATTTAGTTTTTGANGTC 1134

Qy   223 -------------------TyrCysArgArgGlyValIleGlyIlePheGlyProPheTrpGlu 235
                             :::|||
Db  1133 AGATTTCTTACCTAATCCCTACTATGTRAGAGGAATTCCGTCCCATTTACTGGTTTAGNATGAG 1074

Qy   236 HisMetLeuGlyTyr---------TrpArgGluSerLeuLysArgProGlyLysValPhe 252
          |||                      :::||
Db  1073 CCAGTGTACAARTTACGTTATGAAGTGGAAAGAAACC-         CAAATATTT 1029

Qy   253 PheLeuArgGlyTyrGluAspProLeuLysAspIleGluThrArgAsnLeuLysArgLeuAlaThr 272
          |||                                   |||
Db  1028 TTTGATAAATTAACAGATTTA-                CAAATATTT 996

Qy   273 PheLeuGluLeuProPheThrGluGlyGluGlyArgLysGlyValValLysAlaIle 291
          :::|||                                 :::||
Db   995 -ATTCCTGGCTACAAAAAAGAAGGTAAATCGCAATTGGTTATTGCTATA 948
```

; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 55045
; LENGTH: 1316
; TYPE: DNA
; ORGANISM: Bovine
; US-10-750-185-55045

Alignment Scores:
Pred. No.:   17.2
Score: 85.50
Percent Similarity: 36.65%
Best Local Similarity: 24.84%
Query Match: 4.50%
DB: 6

US-10-750-185-55045

Length: 1316
Matches: 40
Conservative: 19
Mismatches: 50
Indels: 52
Gaps: 6

US-10-019-931-3 (1-359) x US-10-750-185-55045 (1-1316)

Qy 116 AsnArgHisArgPheAspProValAlaSerThrAsnHisProLeuPheThrSerAsn 135
         ||| ||||||||::: |||
Db 846 AATTTCTATTTGTTTGACCCGCTCCCACGTCCACAGCTCACTCAGCTCACCCAGTAAAC 905

Qy 136 ProHisAspLeuValProPhePheGluTyrLysLeuTyrAlaAsnGlyAspValProAsp 155
         |||||||| |||||||||
Db 906 CCTGAGGGGTCTGTTTCACTTTTTC---   CTTCCC--- 935

Qy 156 LeuSerGlyLeuAlaSerProArgThrPheAlaThrHisLeuProPheGlySerLeuLys 175
                 |||||||| |||||||
Db 936 ---   ACACAAGCATCACACTGCACACAGTGTGGCCCTTTCCAAA 971

Qy 176 GluThrIleGluLysProGlyValLysValLysTyrLeuCysArgAsnProPheAspThr 195
         :::|||||||
Db 972 CCACCACTGGAAACTTCACTGAAATTGATCATATAT---   1007

Qy 196 PheIleSerSerTrpHisTyrThrAsnAsnIleLysSerGluSerValSer-ProValLe 215
                                                       |||||:||||
Db 1008 ---   CAGGCCAATCTT 1019

Qy 215 uLeuAspGlnAlaPheAspLeuTyr---CysArgGlyValIleGlyPheGlyProPheThr 234
         |||::: ||:||
Db 1020 GATGCCAGAAACCTCTGACATTGTCCCTGCCAGGAATGACAACCCTCACTCGTGCCAAGCTG 1079

Qy 234 pGluHisMetLeuGlyTyrTrpArgGluSerLeuLysArgProGluLysValPhePheLe 254
         ||| |||:||||
Db 1080 GACTCATTGCCCTGGGC---   AAGTGAAGGAGGACAAAATTTCACTTC 1121

Qy 254 uArgTyr-----GluAspLeuLysAspPheIleGluThrAsnLeuLysArgLeuAlaTh 272
         :::   :::||||||||||||||||:::||||||||||
Db 1122 ACTCCATTCGGAGCAACAAATTCAGGACACAATTGAAACACTAGCCTTAAGCGTTAAGACATC 1181

Qy 272 r 272
         :
Db 1182 T 1182

RESULT 10
US-11-121-086-85/c
; Sequence 85, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
;

; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
;
; SEQ ID NO 85
; LENGTH: 171423
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-121-086-85

Alignment Scores:
Pred. No.:   3.78e+04
Score: 85.00
Percent Similarity: 34.69%
Best Local Similarity: 19.90%
Query Match: 4.48%
DB: 7

US-10-019-931-3 (1-359) x US-11-121-086-85 (1-171423)

Qy   20 HisLysLeuGluLeuLeuLysGluGly---   LysThrArgAspValProLys 35
         :::||||||||| |||||:::
Db 21819 TACAAACTGAAAAACTTAGAAGAGATGGATACATTCCTGAACACATACTCTGTCCAAGA 21760

Db 21819 Length: 171423
         Matches: 78
         Conservative: 58
         Mismatches: 143
         Indels: 114
         Gaps: 16

Qy   36 AlaGluGluGluAspGluGly---   LeuSerCysGluPheGlnGlu 48
         ::: ||||| |||
Db 21759 CTGACCAGAGAAAAAACTGATTTCCTGAACAGACCAATATGAGCTCTGAAATTGAATCA 21700

Qy   49 MetLeuAspSerLeuProLysGluArg---   GlyTrpProArgThrArgTyr 63
         :::||| |||:::
Db 21699 GTAGTAAATAGCCTACCAACCAAAAAAGTCCAGTACCTGATGGATTAACAGCTAAATTC 21640

Qy   64 LeuTyrLeuPheGlnGlyPheTrpCysGlnAlaLysGluIleGlnAlaIleIleMetSerPhe 83
         :::|||||||:::
Db 21639 TACCCAGATGTACAAA---   AGAGAGCTACTACCATTCCTA--- 21604

Qy   84 GlnLysHisPheGlnSerLeuGlnAsn---   AspVal 94
         ||||||||| ||| :::
Db 21603 CAGAAACTATTCCAAAAAAATTGAAGAGAAGGGACTCCTCCCCAACTCATTCTGTGAGGCC 21544

Qy   95 ValLeuAlaThrIleProLysSerGly---   ThrThrTrpLeuLysAlaLeuThrPhe 112
         :::||| |||:::
Db 21543 ATCATCATTCTGATACCAAAAACCTGGCAGACACAACAAAAAAAGGAAAACTATAGGTC 21484

Qy 113 ThrIleLeuAsnArgHisArgPheAspProValAlaSerSerThrAsnHisPro--- 130
         :::||||||| :::|||
Db 21483 AATATCCTTAATGACACATHCATTGGGAAAACCCTCAATAAAATACTTGCAAACCAAACCA 21424

Qy 131 ---   LeuPheThrSerAsnProHisAspLeuValPro--- 141
                                |||
Db 21423 TCAAATTAGCACATCAAAAAGCTAATTCACCACAATCAAGGAGGCTTCATCCCTGGGATC 21364

Qy 142 ---   PhePheGluTyrLysLeuTyrAlaAsn---   Gly 151
         :::||| :::
Db 21363 CAAAGTTGGTTCAACATCACACAAAATATATAAGGTGATTAATCACATAAAACGAACTAAA 21304

Qy 152 AspValProAspLeuSerGlyLeuAlaSerProArgThrPheAlaThrHisLeuProPhe 171
         ||| :::
Db 21303 GACAAAACCCACATGATTATCTCAATAATTCAGAAAAGGCTTTCAATAAAATTCAACATC 21244

Qy 172 GlySerLeuLysGluThrIleGluLysProGlyValLysValLysTyrLeu--- 188
         |||||||||:::   :::
Db 21243 AATTCATTGAAAA--ACTCTCAATAAAACTAGGTATTGAACGAGTATATCTCAAAATAATA 21186

Qy 189 ---   CysArgAsnProPheAspThrPheIleSerSerTrpHisTyrThrAsnAsn 205
         ::: |||
Db 21185 AGAGCCATCTCTGTGACAAACCCGCAGCCAACTTTATACTGAATGGGCAAAAATCTGGAGCA 21126

```
RESULT 11
US-11-117-187-186
; Sequence 186, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
;  APPLICANT: PREUSS, DAPHNE
;  APPLICANT: COPENHAVER, GREGORY
;  TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
;  FILE REFERENCE: ARCD.309US
;  CURRENT APPLICATION NUMBER: US/11/117,187
;  CURRENT FILING DATE: 2005-04-28
;  PRIOR APPLICATION NUMBER: US/09/531,120
;  PRIOR FILING DATE: 2000-03-17
;  PRIOR APPLICATION NUMBER: 60/125,219
;  PRIOR FILING DATE: 1999-03-18
;  NUMBER OF SEQ ID NOS: 212
;  SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 186
;  LENGTH: 67088
;  TYPE: DNA
;  ORGANISM: Arabidopsis thaliana
;  FEATURE:
;  NAME/KEY: modified_base
;  LOCATION: (27227)..(64728)
;  OTHER INFORMATION: N = A, C, G, or T/U
US-11-117-187-186

  Query Match:     4.45%

US-10-019-931-3 (1-359) x US-11-117-187-186 (1-67088)

Alignment Scores:
Score:              1.01e+04     Length:              67088
Pred. No.:               3
Percent Similarity:  31.00%      Matches:                94
Best Local Similarity: 19.96%    Conservative:           52
                                 Mismatches:            159
Query Match:         4.45%       Indels:                167
DB:                      7       Gaps:                   19
```

```
Qy    3 ThrSerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHisLysLeu 22
          :::  ||||||||||
Db 18187 ACATTCTCCGTTCACGTCTCTCACTCGACCAAAACATCACTCTCTCACCGCCTC 18246
```

```
Qy  264 GluThrAsnLeuLysArgLeuAlaThrPheLeuGlu-------------------------- 275
        ||||||||||:::||||||||||||::||
Db 19170 TTGAACAGGCCTAAATGAAGTCCCAGTGGAGGAGTACATCAGAGACTTTTCGAGCTGAACGAG 19229
```

```
Qy  259 LeuLysAspAspIle----------------- 263
        ||||||||||
Db 19110 ATCGACGATGATAT... 19169
```

```
Db  19230  TTCTGGGGAACGAGAGTACCCCTATTATCAGACTTTAGCCCAGCTGGGGCTTACCGGAGGAC  19289
                  |||:::                                        :::          |||:::
Qy    288  ValLysAlaIleAlaGluLeuCysSerPheGluAsnLeuLysLysLeuGluValAsnLys  307

Db  19290  GTACAACACCTATTCGAGAAGTGTCATCTCGAGACACTGACTTCTTACCCGTAGTCGCT  19349
Qy    308  SerAsnLysSerIleLeuAsnAsnArgPheLeuAsnArgLysGlyGlyGluValSer  327

Db  19350  TACAAGAAGGAAATAAGAGAGTTT-------------------------------  19373
                 |||:::    |||
Qy    328  AspTrpValAsnTyrLeuSerProSerGlnValGlyArgLeuSerAlaLeuValAlaAspAsp  347

Db  19374  ------------------CTCTCCACTCGTCCAAGTCGAGAGATGTATCAGGACTTACAGCAGAT  19418
                                          |||:::|||||||
Qy    348  LysLeuLeuGlyLysGlyIleLeuThrPhe  356

Db  19419  GAGCTGGAGAGTGAAGGGGTTGGGGTTC  19445
```

```
Alignment Scores:
Pred. No.:
Score:                    4.41
Percent Similarity:       84.00          Length:        420
Best Local Similarity:    32.34%         Matches:       44
Query Match:              21.89%         Conservative:  21
DB:                        4.43%         Mismatches:    60
                          6              Indels:        76
                                         Gaps:          8

US-10-019-931-3 (1-359) x US-10-821-234-822 (1-420)

Qy    156  LeuSerGlyLeuAlaSerProArgThrPheAlaThrHisLeuProPheGlySerLeuLys  175
                  :::|||                   |||:::          ||||||           |||:::
Db      1  ATGTCATCACTACCCGTACCATACACACGCGTTTCCTTGCCTTGTTGGTTCG-----  54

Qy    176  GluThrIleGluLysProGlyValLysValTyrLeuCysArgAsnProPheAspThr  195
                  |||                            |||
Db     55  --------------TGCGTGATAATCACAGGGACGACACGGATCCTCTCACT  87

Qy    196  PheIleSerSerTrpHisTyrThrAsnAsnIleLysSerGluSerValSerProValLeu  215
                  :::                                          |||:::
Db     88  TTTGTCAAGGACACCACAGCGTGGAGGTGAATTTCTACACTGG------  129

Qy    216  LeuAspGlnAlaPheAspLeuTyrCysArgGlyValIleGlyPheProPheTrpGlu  235
                  ||||||          |||:::                   |||:::
Db    130  ATGGATGAGGACTCAGATATTGCTTCCAATTCCGACTGCACTTTGGT------  177

Qy    236  His-------------  ------MetLeuGlyTyrTrpArgGlySerLeuLysArgPro  248
                  |||                          :::
Db    178  CATCCTGCAATCATGAACAGTTGTGTTTGGCATATGGAGATATGAG------  225

Qy    249  GluLysGlyValPhePheLeuAsnArgTyrGluAspLeuLysLysAspAspIleGluThrAsnLeuLys  268
                  |||                                               :::
Db    226  GAGAAAATGCTACTATTACCCTTTGAAGATGGCAAA------  261
```

```
Qy    269  ArgLeuAlaThrPheLeuGluLeuProPheThrGluGluGluLeuArgLysGlyValVal  288
                  :::|||                    |||:::          |||||||
Db    262  ------------------CCATTT------  267

Qy    289  LysAlaIleAlaGluLeuCysSerPheGluAsnLeuLysLysLeuGluValAsnLysSer  308
                  |||:::                   |||:::          :::
Db    268  ------------------GAGCTGTGCATCTATGTCCGTCACAAGAATACAAAGGTAATGGTAAAT  315

Qy    309  AsnLysSerIleLeuAsnAsnArgPheLeuAsnArgLysGlyGlyGluValSerAspAsp  328
                  |||:::                   |||:::
Db    316  GGCCAACGCATTTACAACTTTGCCCATCGATTC------  348

Qy    329  TrpValAsnTyrLeuSerProSerGlnValGlyArgLeuSerAlaLeuValAlaAspAsp  348
                  |||:::|||:::          :::|||
Db    349  ------------------CCGCCAGCATCTGTGAAGATCCTGCAAGTCTTCAGAGATATCTCC  393

Qy    349  Leu  349
                  |||
Db    394  CTG  396
```

```
Alignment Scores:
Pred. No.:
Score:                    3.73e+04
Percent Similarity:       84.00          Length:        142303
Best Local Similarity:    33.41%         Matches:       84
Query Match:              20.49%         Conservative:  53
DB:                        4.43%         Mismatches:    114
                          7              Indels:        159
                                         Gaps:          22

US-10-019-931-3 (1-359) x US-11-121-086-42 (1-142303)

Qy      4  SerSerMetLysSerIleProMetAlaIleProSerPheSerMetCysHisLysLeuGlu  23
                  :::|||||||                        :::
Db  95742  AGTTCAGCCCAGACACAAGTTAAAATGTTCTTCCACAGAACAGTCCTGCCTCAGAAATGAA  95683

Qy     24  LeuLeuLysGluGlyLysThrArgAspValProLysAla-------GluGluAsp  39
                  :::     :::                                          |||:::
Db  95682  TGAACCAGCCAGGCACGGTGCCGTGGAGGCCCCTGTAATCCCAGAATTTGGAGGCCGAG  95623

Qy     40  GluGlyLeuSerCysGlyGlnGlyMetLeuAspSerLeuProLysGlnArgGlyTrp  59
                  :::          |||                        |||:::          :::
Db  95622  ATGGGAGGAGGATCACTTGAGGTCAGGAGTTTGAGACCAGCCTGGCCAACA------  95572

Qy     60  Arg------------ThrArgTyrLeuTyrLeuPheGlnGlyPheTrpCys--  72
                  :::                         :::
Db  95571  CAAAACTCTGTCTCTACTAAAAAATACAAAAATTAGCTGGGCGTGGTGGCATGGCCTGTA  95512

Qy     73  ----------------GlnAlaLysGlu--------  77
                  :::|||:::|||
Db  95511  ATTCCAGCTACTCAGGAGGCTGAGGAGGAGGATTGCTTGAACCCAGGAGGCAGGGGTTG  95452
```

RESULT 14

```
Qy   78 GlnAlaIleMetSerPheGlnLysHisPheGlnSerLeu---GluAsnAspValValLeu 96
                |||                                   :::   |||
Db 95451 CAGTGAGCCGAGATGTACAAATGACACTCCAGCGCGGGGTGAAAGAGGAGACCCTATCTCA 95392

Qy   97 AlaThrIleProLysSerGly---ThrThrTrpLeuHisAlaLeuThrPheThrIleLeu 115
           |||                   :::              ||| |||
Db 95391 AAAGAAATAAAAAGAAATGGAATGAACCGTTGGCGTGCAGTACTGCACC------CTA 95341

Qy  116 AsnArgHisArgPheAspProVal-AlaSerSerThrAsnHisProLeuPheThrSerAs 135
           |||                       :::  |||   |||||||
Db 95340 CGGCCCTTTTTATTGGATGTGAGTGTTCCTATCCTTAAGGCACACTTTATTTACTAGAAC 95281

Qy  135 nProHisAspLeuValProPhePheGluTyrLysLeuTyrAlaAsnGlyAspValProAs 155
                |||                     |||||            :::
Db 95280 AAGATTAGATAAAGAACCATTA------AAAAATTCTTCTGTATTGGGAGTCTCGAGGTG 95230

Qy  155 pLeuSerGlyLeuAlaSerProArgThrPheAlaThrHis------ 168
               :::                       |||
Db 95229 GGTGGGGGGGATTGCTTGAGCCCAGGAGTATGAGAACACAGCTGGGGACCC 95170

Qy  168 uCysArgAsnProPhe------AspThrPheIleSerSerTrpHis- 201
              :::                   |||
Db 95169 C---AGACTACCACTCCCTGCGAAACTTCCCTGCTCACCACCACTGCTTTTTCTGAC 94993

Qy  169 -LeuProPheGlySerLeuLysGluThrIleGluLysProGlyValLysValValTyrIle 188
               :::                            |||  :::
Db 95109 GTTCCCTAGAGACACTCTGACTGCGTGGGTTGAAGGATGAGGTAGAAGGTGGCAGAGAT 95050

Qy  202 -----TyrThr--- 203
            |||
Db 94992 TGGACTTGCTTGACAGGTGGTTAGTTCAGCATAATAAGTTCTTAGCTCAGTTAACTACACTC 94933

Qy  204 ----AsnAsnIleLysSerGluSerValSerProValLeuLeuAsp- 218
                :::
Db 94932 GGGGCTCAGTTAACTACACTCAGGGTTGGTCTTATTTCTGTGGGGGGGGGGAGTGAGAC 94873

Qy  218 nAlaPheAspLeuTyrCys------ArgGlyVal--Il 228
              |||                   :::||||
Db 94872 AGAAGATAGCCTAAGATTTTCCAGGTTCTCCCCTATATTGTCTGATGGCCAGAAAATTCCAC 94813

Qy  228 eGlyPheGlyProPhe------TrpGluHisMetLeuGlyTyrTrpAr 242
              |||||                 ||||
Db 94812 AGCATTTCACCTCTTGGTCACATTGAAAGTAAACATTCTCCCCAAGAAGGATTAGTTGGAG 94753

Qy  242 gGluSerLeuLysArgProGlyLysProGlyLysGlu----- 257
           ||||
Db 94752 CGGGCTTTGGAGTTTTCATTACAGGGGAGTGAGGGGTGGGAACAT 94710

Qy  258 -----AspLeuLysAspIleGluThrAsnLeuLysArgLeuAlaTh 272
               :::
Db 94709 ----AAGGAGCCTGGGCAAAATGTGGAGGATAAGGAAAACAGAGCCAGAGAGGG 94660

Qy  272 rPheLeuLeuProPheThrGluGlu 281
              :::||||||||
Db 94659 GCTTTCTTTACTGTGAAATGTTTGTGATGACATTGACGAGAC------AC 94618

Qy  281 ...
Db 94617 ATTCCTAGAC-----TGGACAGAGCAG 94596
```

Alignment Scores:

Pred. No.:
Score: 28.6
Percent Similarity: 83.50
Best Local Similarity: 40.24%
Query Match: 26.04%
                            4.40%
                            6

Length:	1281
Matches:	44
Conservative:	24
Mismatches:	58
Indels:	43
Gaps:	10

US-10-019-931-3 (1-359) x US-10-750-185-31307 (1-1281)

```
Qy   15 SerPheSerMetCysHisLysHisLeuGluLeuLeuLys----------------- 26
           ||||||||            |||:::
Db  969 AGCAGGAGCCCGTCGCCATCCTCTCAACTTACTAGGTGGCTCAGTGATAAAGAATCTGCC 910

Qy   27 -----GluGlyLysThrArgAspValProLysAlaGluGluAspGlu------ 41
               :::
Db  909 TGCAATGTCAGGAGACTCAGGTTCGATCCCTGGGTCGGGAAGATCCCCTGGGACGAGGAAT 850

Qy   42 ----LeuSerCysGluPheGlnGluMetLeuAspSerLeuProLysGluArgGly 58
              |||
Db  849 GGCAACCTACTCCACGTGTTCTTGCCTGGAGAATCTC-----CTGACAGAGGAGCCTGG 796

Qy   59 Trp----------ArgThrArgTyrLeuTyrPheGlnGly 69
          |||
Db  795 TGGGCTACAGTCCGGCGTGGGTCACAAAGAGTCGGACTGAGCATGTAGGTGTCTTTCAATAG 736

Qy   70 PheTrpCysGlnAlaLysGluIleGluGlnAla------IleMetSerPheGlnLysHisPhe 87
          |||                                :::
Db  735 GTCTGGGCCAAGACGTGAGATGTCGTTCCTTTTCCTCTGCCATTTAGGTGACACTTC 676

Qy   88 ---GlnSerLeuGluAsnAspValValLeuAlaThrIleProLysSerGlyThrThrTrp 106
             |||                            :::||
Db  675 TGACTGCTTCTCGGCCTCCTCGATCGTCCTTTTA-----ATTCCAGCCACCCAGCCCAGTCT 622

Qy  107 LeuLysAlaLeuThrPheThrIleLeuAsnArgHisArgPheAspProValAlaSerSer 126
          |||||                            :::
Db  621 CTGAACCATCTCACTTCTGACAATGTTCCCTTGACAT------CCAGTTTCTTCCAGA 571

Qy  127 Thr------AsnHisProLeuPheThrSerAsn------ 138
          :::
Db  570 TCCATCTGCAGACATCCTTGGTTCCTGATTTCCTTGGTTAATTTCTGATGACCTTCTCGCC 511

Qy  139 LeuValProPhePheGluTyrLysLeu 147
          |||||||||            :::||||
Db  510 TCTGTCCCCTTTCCTTGATCACAAACTT 484
```

RESULT 15

```
; APPLICANT: Stoop, Johan M.
; TITLE OF INVENTION: Plastidic Phosphoglucomutase Genes
; FILE REFERENCE: BB1451 USCIP
; CURRENT APPLICATION NUMBER: US/11/094,586
; CURRENT FILING DATE: 2005-03-30
; PRIOR APPLICATION NUMBER: US 09/906,209
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: US 60/218,712
; PRIOR FILING DATE: 2000-07-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Microsoft Office 97 & PatentIn Version 3.3
;
; SEQ ID NO 1
;   LENGTH: 2112
;   TYPE: DNA
;   ORGANISM: Typha latifolia
US-11-094-586-1
```

```
Alignment Scores:
Pred. No.;
Score:                      81.5
Percent Similarity:         82.50      Length:        2112
Percent Identity:           30.99%     Matches:         55
Best Local Similarity:      19.37%     Conservative:    33
Query Match:                4.35%      Mismatches:      87
                            7          Indels:         109
                                       Gaps:            12
```

US-10-019-931-3 (1-359) x US-11-094-586-1 (1-2112)

```
Qy   114 IleLeuAsnArgHisArgPheAspProValAlaSerSerThrAsnHisProLeuPheThr 133
             ||||||||                  ||:::
Db  1108 ATTCTTGGAAGACGTTTCTTT--------GTTACACCATCAGATTCGTTGCAATGANTGCC 1161

Qy   134 SerAsnProHisAspLeuValProPhePheGluTyrLysLeuTyrAlaAsnGlyAspVal 153
             :::|||||  :::||||||:::  :::
Db  1162 GCTAATGCACAGGCGGGCTATTCCTTATTCCAA---------------------------- 1194

Qy   154 ProAspLeuSerGlyLeuAlaSerProArgThrPheAlaThrHisLeuPro---PheGly 172
             ||| :::                ||:::
Db  1195 ------GCTGGTCCCAAAGGACTTGCTAGGTCTTATGCCAACAAGCGGGT 1236

Qy   173 SerLeuLysGluThrIleGlyLysProGlyValLysValThrTyrLeuCysArgAsnPro 192
             ::::::::::::|||
Db  1237 GCTCTTGATCGTGTGAGCCGAAAAATTGAACCTT-------------------CCA 1272

Qy   193 PheAspThrPheIleSerSerTrpHisTyrThrAsnAsnIleLysSerGluSerValSer 212
             ||||:::
Db  1273 TTCTTTGAGGTTCCAACTGGTTGGAAGTTTTTTGGAAATCTGATGGATGCTGGGAAGTTG 1332

Qy   213 ProValLeuLeuAspGlnAlaPheAspLeuTyrCysArgGlyValIleGlyPheGlyPro 232
             :::::::::::||
Db  1333 TCCATCTGTGGGGAGGAAAGTTTT-------------------GGCACAGGTTCT 1368

Qy   233 ------------------------PheTrpGluHisMetLeuGly 239
                                        :::
Db  1369 GATCACATCCGGGAGGAAGGATGGCATCTGGGCGTGTTTGGCTTGGCTTTCCATAATTGCG 1428

Qy   240 TyrTrpArgGluSerLeuLysArgProGlyLysVal- 251
             |||              :::
Db  1429 TACAGAAACAAGGACAAAAAGAGATTGGAGAGAAATTAGTCTTCTGTTGAAGATATTGCTAAG 1488

Qy   252 ----------------------PhePheLeuArgTyrGluAspLeuLysAspAsp 262
                                       ||||  :::
Db  1489 GAGCACTGGGCAAAATAGGCAGGAACTTCTTTTCTCGATATGAT-------------- 1533

Qy   263 IleGluThrAsnLeuLysArgLeuAlaThrPheLeuGluLeuProPheThrGluGluGlu 282
                                                       ||| :::
Db  1534 ------------------------------------------TACGAAGAATGC 1545

Qy   283 GluArgLysGlyValValLysValAlaIleAlaGluLeuCysSerPheGluAsnLeuLysLys 302
             :::|||
Db  1546 GAATCGGAAGGAGGCAAAATAAAATGATGCAGCACCTTAGGGACTTT-------------- 1590

Qy   303 LeuGluValAsnLysSerAsnLysSerIleGlyAsnPheGluAsnArgPheLeuPheArg 322
             ::::|||
Db  1591 ATCTGACAAGCAAGCACGCCTGAGAACAAATATGGAAAATTATACTCTTCAATTT------ 1641

Qy   323 LysGlyGluValSerAspTrpValAsnTyrLeuSerProSerGlnValGluArgLeuSer 342
             :::|||||               :::
Db  1642 ------TCAGATGACTTTCCTACACTGACCCT-------------------GTAGAC 1674

Qy   343 AlaLeuValAspAspLysLeuGly------------------GlySerGlyLeuThr 355
                                                        :::
Db  1675 GGCAGGTGTAGCATCCAAGCAAGGGCTACGATTTGTTTTCACAGATGGATCAAGGGGTTATC 1734

Qy   356 PheArgLeuSer 359
             ::::|||||||
Db  1735 TATCGTCTCTCG 1746
```

Search completed: December 23, 2005, 22:08:33
Job time : 489 secs